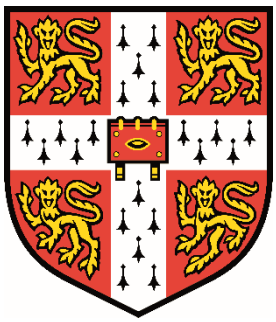
# Mechanisms of Oct4 in the entry to, maintenance of, and exit from pluripotency

Lawrence E. Bates

This dissertation is submitted for the degree of Doctor of Philosophy

Supervised by José Silva, PhD

St Edmund's College
University of Cambridge



June 2019

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as specified in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as specified in the text.

It does not exceed the prescribed word limit for the relevant Degree Committee.

Lawrence Bates

September 2018

*To my niece, Betsy*

## Acknowledgements

I would like to thank José for his support and direction. In addition to offering me a position in his group, he has offered me advice, reassurance, and optimism. His instinct and insight have been invaluable, and his confidence in me even when experiments failed buoyed my spirits. The balance between mentorship and independence that he offered has helped me to grow as a scientist.

I would like to thank all the members of the Silva lab, it has been a pleasure working with you. Aliaksandra Radzisheuskaya, Rodrigo Santos, Morya Lawrence, Hannah Stuart and Yael Costa welcomed me into the group when I was a new research assistant. It was inspirational to work alongside confident and skilled scientists. It was a pleasure to work alongside Charlotte Handford, Elsa Sousa, Katie Tremble, Kate Maskalenka, Chibeza Agley, and Sergey Gladkov. I would also like to thank the undergraduate and masters students I worked with; Aran Shaunak, Saeed Kayhanian and Mariana Alves contributed both to the progression of this work and to the friendly atmosphere of our lab.

I'm grateful to the PIs of the Wellcome - Medical Research Council Stem Cell Institute for giving me the opportunity to study here, and to the MRC for funding my position. I'd also like to thank everyone that helps to keep the institute running, and I'd particularly like to thank Sally Lees, Peter Humphreys, and Maike Paramor and Michael Barber. Without their expertise, none of this would be possible.

I would like to thank Jenny Nichols, Austin Smith, Kevin Chalut, Brian Hendrich, and others for their insightful comments, advice and suggestions.

There are many other PhD students and post-docs from the institute and the wider university, too many to name individually, who I would like to thank for their help, stimulating conversations, and the occasional therapy session in the pub. Members of the Smith, Nichols and Chalut labs in particular had birthdays with an abnormal frequency.

Amanda Andersson-Rolf and Hannah Stuart were both wonderfully supportive both in the lab and out, and I'm grateful to you both for your love and encouragement. Away from the institute, Anna Blackwell and Nick Garidis offered me help and support when I needed it, without you I would never have made it this far.

Finally, I'd like to thank my family for their ceaseless support, encouragement and love, and complete lack of understanding of what I do.

## Abstract

Pluripotency is defined as the capacity to give rise to all cell types of the embryo proper. It arises in the early mammalian embryo but is lost after a short period of time as cells differentiate and become committed to different lineages. Prior to implantation, mouse epiblast cells enter the pluripotent naïve state, which can be captured *in vitro* in the form of embryonic stem cells. These cells are characterized by a capacity for indefinite self-renewal, and the ability to re-enter normal development upon being returned to the naïve epiblast. A complex transcription factor network promotes this state. Overexpression of one of many of these factors leads to stabilisation of the naïve state, with enhanced self-renewal and reduced spontaneous differentiation. However, the transcription factor Oct4 must be maintained within a tight window of expression; depletion results in extraembryonic differentiation, while overexpression also results in exit from pluripotency.

Oct4 was identified as a protein expressed in the early embryo and in germ cells, and was subsequently discovered to be essential for the establishment of the naïve epiblast. *In vitro* studies determined that loss of Oct4 in ESCs induced trophoblast differentiation. Meanwhile, overexpression of Oct4 led to differentiation of ESCs, and constitutive expression of Oct4 was not sufficient to replace any of the extrinsic factors required for ESC self-renewal. Despite these dramatic phenotypes, the essential role of Oct4 remains unclear, further complicated by the finding that a reduced level of Oct4 promotes self-renewal at the expense of differentiation capacity.

In this work, I generated a novel Oct4 fusion protein capable of rapid inducible degradation in order to study the immediate responses to removal of Oct4. This system utilizes the auxin responsive degradation domain of the *Arabidopsis thaliana* IAA17 protein to recruit a transgenic F-box protein Tir1 on addition of the small molecule auxin to the culture medium. Subsequent ubiquitination by the endogenous SCF complex leads to rapid proteolytic degradation of the Oct4 fusion protein, resulting in loss of detectable protein in as little as two hours. This system allows the study of immediate responses to loss of Oct4 in contrast to conventional depletion systems in which Oct4 levels decay over a protracted period making it difficult to disentangle direct and indirect effects.

I established that several pluripotency-associated genes require Oct4 for their transcription. RNA levels of these factors decrease rapidly on depletion of Oct4, prior to significant changes in the expression of other key pluripotency factors such as Nanog and before protein levels of other factors

can change dramatically. Furthermore, I established that the presence of Oct4 antagonises chromatin binding by Nanog, revealing a possible mechanism by which increased Oct4 levels are detrimental to the naïve state. Together, these findings may be sufficient to explain the simultaneous requirement for, and antagonistic activity of, Oct4 in naïve pluripotent cells.

I also found that an ESC level of Oct4 facilitates cell identity transitions. Cells constitutively expressing Oct4 and differentiated *in vivo* could be reverted to naïve pluripotency *in vitro* through the application of defined naïve pluripotency growth conditions. Additionally, in the course of these experiments I examined the phenotypic abnormalities that occur in mouse embryonic development under continuous expression of Oct4. In keeping with previous work, we observed abnormalities in limb development and in the skin. We also observed exencephaly in a number of embryos.

In conventional exit from pluripotency, female cells must inactive an X chromosome in order to balance X-linked gene expression between males and females. This is achieved via expression of Xist from a single X chromosome, where it orchestrates chromosome-wide silencing. I show that Oct4 plays an important role in the regulation of Xist during the exit from pluripotency. I find that, in the absence of Oct4 or in extraembryonic differentiation induced by other methods, such as overexpression of Cdx2 or Gata6, downregulation of Xist does not occur in male cells resulting in persistent high levels of Xist expression. Normally, Oct4 expression persists after the downregulation of most naïve transcription factors during early differentiation. I propose that this allows Oct4 to antagonise expression of the Xist, and thus ensure proper control over X chromosome inactivation.

Together this work focuses on the dual roles of Oct4 in regulating both pluripotency and differentiation. I address the contradictory phenotypes relating to altered expression of Oct4, and establish a unifying theory to explain them. I put forward evidence that Oct4 promotes cell fate transitions by regulating naïve transcription factors. I propose that the environment plays a key role in determining cell identity, while Oct4 acts to maintain plasticity by preventing cells from being trapped within otherwise stable states.

## Contents

## List of Figures

## List of Abbreviations

| | |
|---|---|
| 2i | Combined Chiron and PD03 |
| 4-OHT | Synthetic ER agonist 4-hydroxytamoxifen |
| A-P | Anterior-posterior axis |
| AID | Auxin inducible degron |
| Bsd | Blasticidin S or the blasticidin resistance gene |
| CAG | A synthetic constitutive promoter |
| ChIP | Chromatin immunoprecipitation |
| ChIP-qPCR | ChIP followed by qPCR analysis |
| ChIP-seq | ChIP followed by NGS and downstream analysis |
| Chiron | Pharmacological GSK3 inhibitor CHIR99021 |
| $CreER^{T2}$ | Fusion protein of Cre and G400V/M542A/L544A triple mutant human ER |
| DAPI | 4′,6-diamidino-2-phenylindole, a fluorescent DNA stain |
| DMSO | Dimethyl sulfoxide |
| Dox | Doxycycline, a tetracycline derivative |
| E | Embryonic day (e.g. E4.5 is 4.5 days into embryonic development) |
| EGC | Embryonic germ cell |
| EpiLC | EpiSC-like cell |
| EpiSC | Epiblast stem cell |
| ESC | Embryonic stem cell |
| F | Floxed allele |
| FCS | Fetal calf serum |
| FDR | False discovery rate |
| Floxed | Flanked by LoxP sites |
| GOI | Gene of interest |
| H&E | Haematoxylin and eosin |
| Hyg | Hygromycin B |
| IAA | Indole acetic acid, an auxin plant hormone |
| IF | Immunofluorescence |
| iPSC | Induced pluripotent stem cell |
| MEF | Mouse embryonic fibroblast |
| NGS | Next generation high throughput sequencing |
| NSC | Neural stem cell |
| qPCR | Real time quantitiative PCR |
| PB | piggyBac transposon |
| PCR | Polymerase chain reaction |
| PD03 | Pharmacological MEK1/2 inhibitor PD0325901 |
| Pre-iPSC | Reprogramming intermediate, precursors to iPSCs |
| Puro | Puromycin |
| RNA-seq | NGS of RNA |
| RT-qPCR | Reverse transcriptase qPCR |
| rtTA | Reverse-tTA |
| SL | Serum-based media supplemented with LIF |

| | |
|---|---|
| TRE | Tetracycline response element, a series of TetO elements followed by a minimal CMV promoter; active when tTA or rtTA is bound. |
| TSC | Trophoblast stem cell |
| tTA | Tetracycline responsive transactivator, a fusion protein of TetR and VP16 |
| XCI | X chromosome inactivation |
| XIC | X inactivation centre |
| ZPA | Zone of polarizing activity |

## List of Tables



## List of Equations



## List of Lists

## List of Appendices

# Chapter 1: Introduction

## 1.1. Pluripotency and early mouse development

Development is often seen as a gradual restriction in fate potential. Embryogenesis begins with a fertilized oocyte, a single zygote with the potential to form every cell of a new organism and all of the extraembryonic tissue required to support embryonic growth. Over time, groups of cells differentiate, heading down paths limited to extraembryonic identities, while others lose this capacity but maintain the potential to contribute to the embryo proper; these are pluripotent cells. As development continues cells are segregated into major cell types, the three germ layers and the germ line, where they continue to differentiate, gradually losing their potency as they become specialized and adopt specific roles around the body. This is wonderfully captured in the concept of Waddington's landscape (Waddington, 1957), a mountain face with peaks and troughs that direct a rolling ball to one of several end points; a model of growing epigenetic constraints that direct cells to their identities and stably hold them there.

### 1.1.1. The early mouse embryo

Mammalian development has been best characterized in mouse. This is a model system adopted for ease of handling and rapid reproductive cycle, and reinforced by the development of techniques that allow us to probe the intricacies of development in detail. The earliest stages of mouse development are characterized by cleavage divisions of the totipotent zygote (E0, Figure 1.1) to generate a morula (~E2.0, Figure 1.1).

Figure 1.1

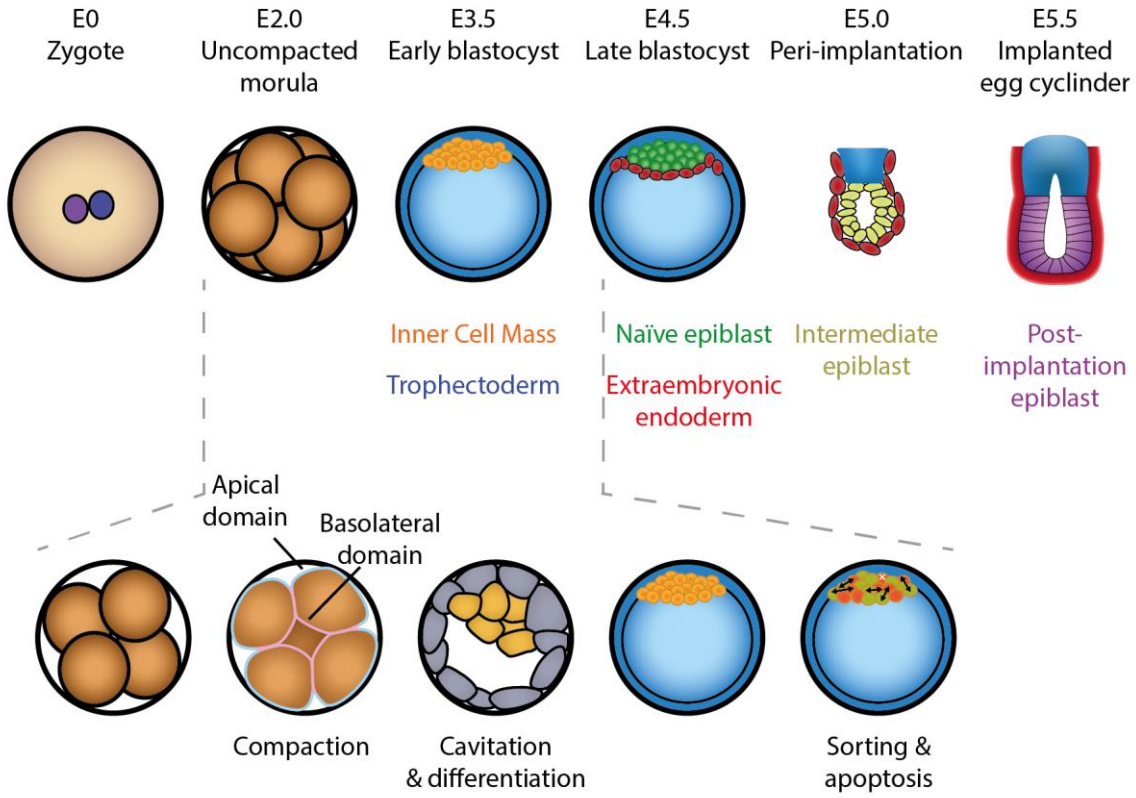



Figure 1.1 Mouse embryogenesis from fertilization to early post-implantation
An overview of early mouse embryogenesis. A fertilized egg (E0) contains the paternal and maternal pronuclei. As this zygote undergoes mitosis the maternal and paternal chromosomes intermingle. After three rounds of cell division (~E2.0), the embryo has reached the state of an uncompacted morula. At around the 8-cell stage, cells convert from being largely unpolarised to possessing distinct apical-basal polarity, and become strongly adherent. This results in a compacted morula in which it is difficult to observe cell boundaries. As these cells continue to divide, some cells are located on the outside of the embryo while others are located inside the morula. This distinction plays an integral role in the differentiation of the outside cells to trophectoderm and the inside cells form the inner cell mass and lose their polarity. A cavity forms within the embryo, with cells of the inner cell mass clustered at one side, firmly attached to the trophectoderm (~E3.5). Differences in expression of factors such as Nanog and Gata6 result in a salt-and-pepper mixture of immature epiblast and extraembryonic endoderm cells. A mixture of cell sorting and apoptosis leads to an organised structure with the naïve epiblast in contact with trophectoderm, and extraembryonic endoderm separating the epiblast from the blastocoel (~E4.5). The zona pellucida is shed and the embryo implants (~E5.0), with the trophectoderm expanding and invading the maternal endometrium, and forming the yolk sac together with the extraembryonic endoderm. At the same time, the epiblast undergoes significant changes, forming a monolayered epithelium with a central cavity (~E5.5). Note that mouse ESCs are approximately equivalent to the E4.5 naïve epiblast and EpiSCs are roughly equivalent to the E5.5-E6.5 post-implantation epiblast.

The zygotic genome is activated early in mouse, perhaps in the zygote itself (Aoki et al., 1997; Bouniol et al., 1995); nonetheless, large quantities of maternally derived mRNAs are present to direct these initial divisions (Hamatani et al., 2004). After three cycles of cell division, the eight-cell embryo begins to undergo compaction (Ducibella et al., 1977; Pratt et al., 1982). Cells obtain apical-basal polarity (Ducibella et al., 1977; Vinot et al., 2005) and adhere tightly together (Ducibella and Anderson, 1975; Vestweber et al., 1987) (~E2.5, Figure 1.1, inset), and as cells continue to divide the inner cells become surrounded by basolateral domains (Hirate and Sasaki, 2014). Cells on the outside of the embryo

differentiate towards trophectoderm through the combined activity of YAP/TAZ and the transcription factor Tead4, while inner cells remain undifferentiated as their close contact leads to activation of the Hippo pathway and degradation of YAP/TAZ (Hirate et al., 2013; Nishioka et al., 2009, 2008; Yagi et al., 2007). Consequently, the outer cells upregulate trophoblast factors Cdx2 and Eomes while silencing Oct4 (Nishioka et al., 2009; Niwa et al., 2005), while expression of Oct4, Sox2, Nanog, Pdgfrα and Gata6 become more restricted to the inner cells (Kang et al., 2013; Le Bin et al., 2014; Wicklow et al., 2014), now forming the inner cell mass (ICM) as a blastocoel-filled lumen opens within the embryo (Figure 1.1, inset). At this point, antagonistic regulation between Oct4 and Cdx2 cements this first fate decision (Dietrich and Hiiragi, 2007; Niwa et al., 2005).

As the ICM matures, the cells begin to adopt different fates with mutually exclusive expression of Nanog and Gata6, marking cells of the presumptive naïve epiblast and hypoblast respectively (Schrode et al., 2014; Yamanaka et al., 2010) (~E4.0, Figure 1.1, inset). A mixture of cell sorting and selective apoptosis leads to segregation of these cells (Meilhac et al., 2009; Plusa et al., 2008) such that the hypoblast overlies the naïve epiblast, which is in direct contact with the trophectoderm. At around this time, the zona pellucida is shed and embryos are suitable for implantation. If the previous litter has not yet been weaned, signals from the mother can induce embryonic diapause at this point, arresting development prior to implantation (Mantalenakis and Ketchel, 1966; Weitlauf and Greenwald, 1968). Upon implantation (~E5.0, Figure 1.1), the epiblast undergoes dramatic morphological, epigenetic and transcriptional changes. Cells begin to polarize and generate a single layered epithelium as a central lumen forms which will become the pro-amniotic cavity (Bedzhov and Zernicka-Goetz, 2014; Coucouvanis and Martin, 1995) (~E5.5, Figure 1.1). Shortly after this, the visceral endoderm formed from the differentiating hypoblast gains axial polarity, establishing the distal and subsequently anterior visceral endoderm (Beddington and Robertson, 1999; Takaoka et al., 2011, 2006). Signalling from this anterior-specifying region and the overlying trophectoderm lead to differentiation of the posterior epiblast into the primitive streak (Tam and Behringer, 1997; Williams et al., 2012), characterized by a transition to a multi-layered epithelium. At this site epiblast cells undergo an epithelial to mesenchymal transition and ingress, passing through the streak and segregating into mesoderm and endoderm while the ectoderm remains on the exterior in the process of gastrulation (Lawson and Pedersen, 1987; Quinlan et al., 1995; Tam and Beddington, 1987; Tam and Behringer, 1997). As they become determined in their new identity, these cells lose their ability to

contribute to other germ layers, and hence lose their pluripotency (Beddington et al., 1992; Carey et al., 1995; Kanatsu and Nishikawa, 1996; Tam et al., 1997).

## 1.2. Studying pluripotency *in vitro*

It was predicted that cells of the pre-implantation mouse embryo might be pluripotent when it was demonstrated that small numbers of isolated blastomeres injected into the blastocoel of a recipient embryo could contribute to all germ layers of the subsequent embryo, generating chimeric offspring (Gardner, 1968). Both pre- and post-implantation embryos can generate tumours containing cells from all three germ layers as well as proliferating undifferentiated cells – teratocarcinomas - on transfer to the testes of adult mice (Stevens, 1970). Single carcinoma cells derived from teratocarcinomas can regenerate tumours containing various somatic tissues, demonstrating functional pluripotency at a cellular level (Kleinsmith and Pierce, 1964). These tumours can also form spontaneously from primordial germ cells (Stevens, 1967), particularly in certain inbred mouse lines (Stevens and Hummel, 1957; Stevens and Little, 1954). This led to the capture of pluripotency *in vitro* in the form of embryonic carcinoma cells (ECCs) (Finch and Ephrussi, 1967; Rosenthal et al., 1970). Incredibly, these cells can contribute to normal, non-tumorous mice on microinjection into blastocysts (Papaioannou et al., 1975; Stewart and Mintz, 1981).

### 1.2.1. Establishing embryonic stem cells

By studying conditions that favour the self-renewal rather than differentiation of ECCs, environments favourable for derivation of pluripotent cell cultures direct from mouse blastocysts were developed, yielding embryonic stem cells (ESCs) (Evans and Kaufman, 1981; Martin, 1981).

Figure 1.2

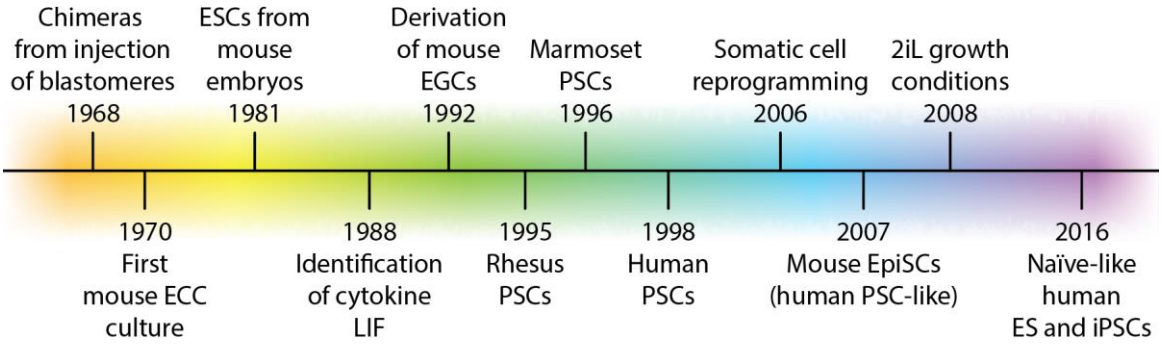



Figure 1.2 Timeline of key discoveries relating to naïve pluripotency in mouse
In 1968 Richard Gardner demonstrated that live chimeras could be born following injection of isolated blastomeres into host blastocysts, implying pluripotency at a cellular level (Gardner, 1968). Soon after, in 1970, the first pluripotent cell lines that display long-term self-renewal without loss of potency were captured *in vitro* by Miriam Rosenthal in the form of embryonic carcinoma cells (ECCs) derived from teratocarinomas (Rosenthal et al., 1970). Following this, embryonic stem cells (ESCs) were independently captured from pre-implantation mouse embryos by Gail Martin and by Martin Evans and Matthew Kaufman in 1981 (Evans and Kaufman, 1981; Martin, 1981). In 1988, the cytokine LIF was identified independently by Austin Smith and Lindsay Williams as a key factor in ESC medium that greatly enhanced the self-renewal of these cells (Smith et al., 1988; Williams et al., 1988). It was independently demonstrated that mouse primordial germ cells could readily convert to pluripotent embryonic germ cells (EGCs) in 1992 by Yasuhisa Matsui and James Resnick (Matsui et al., 1992, 1991; Resnick et al., 1992). Pluripotent stem cells (PSCs) were derived by James Thomson from the preimplantation embryos of rhesus monkeys, marmosets and humans in rapid succession in 1995, 1996 and 1998 respectively by (Thomson et al., 1998, 1996, 1995); due to their pluripotent character and derivation from the preimplantation epiblast, they were termed ESCs. In 2006, Kazutoshi Takahashi demonstrated that it was possible to reprogram somatic cells into induced pluripotent stem cells (iPSCs) through the expression of four transcription factors (Takahashi and Yamanaka, 2006). Keisuke Okita followed up this work and generated chimera-competent iPSCs the following year (Okita et al., 2007). Cell lines closely resembling conventional primate PSCs were derived from post-implantation mouse embryos by both Paul Tesar and Gabrielle Brons in 2007 (Brons et al., 2007; Tesar et al., 2007). In 2009, naïve-specific '2iL' growth conditions that favour the self-renewal of mouse ESCs in a state that closely resembles the preimplantation epiblast were identified by Qi-Long Ying (Ying et al., 2008). More recently, conditions for the propagation of human PSCs in a more naïve-like state have been generated by Thorold Theunissen and Yasuhiro Takashima (Takashima et al., 2014; Theunissen et al., 2014). These 2016 papers were followed up by more extensive molecular characterization by Theunissen and direct capture of naïve-like cells *in vitro* from preimplantation human embryos by Ge Guo respectively (Guo et al., 2016; Theunissen et al., 2016).

On introduction into blastocysts, these cells could re-enter normal development and contribute to all three germ layers plus the germ line (Bradley et al., 1984). These cells possess a robust, highly redundant transcriptional network that maintains their pluripotent character (Niwa, 2018). They are a powerful model system for studying general principles of mammalian cell biology and cell fate transitions since they can be readily expanded to large numbers and can self-renew indefinitely without needing to be immortalized or transformed.

### 1.2.2. The ESC signalling environment

For a long time, mouse ESCs were maintained in cytokine-rich serum-based medium on a feeder layer of mouse embryonic fibroblasts (MEFs) (Evans and Kaufman, 1981), or in ECC-conditioned medium

(Martin, 1981). Investigations into the factors involved in ESC self-renewal in these conditions uncovered the cytokine LIF (Smith et al., 1988; Williams et al., 1988), which enabled propagation of mouse ESCs in the absence of feeders in serum-based media supplemented with LIF (SL). This ligand binds to LIFR with GP130, resulting in activation of the intracellular kinase JAK. JAK phosphorylates STAT3, inducing dimerization leading to nuclear translocation whereupon it acts as a pluripotency-promoting transcription factor (Figure 1.3) (Matsuda et al., 1999; Niwa et al., 1998). It can be shown that the JAK/STAT3 pathway is the major role of this cytokine since Stat3 knockout cells show little response to addition of LIF and Stat3Y705F mutants which cannot be phosphorylated by JAK are not capable of self-renewal in SL conditions (Huang et al., 2014), while constitutively active STAT3 can maintain ESCs in the absence of LIF (Huang et al., 2014; Matsuda et al., 1999).

Further studies found that BMPs contribute to pluripotency and permit self-renewal of ESCs in serum-free media containing LIF (Ying et al., 2003). It appears that BMP signalling via SMAD1 and SMAD5 results in upregulation of Id1 and Id3 which inhibit neural differentiation (Figure 1.3) (Hollnagel et al., 1999; Ying et al., 2003). Overexpression of these Id genes is sufficient for ESC self-renewal in serum-free media containing LIF alone (Ying et al., 2003). Interestingly, conventional BMP-SMAD1/4/5 signalling appears to be entirely disposable for the maintenance of mouse ESCs (Gomes Fernandes et al., 2016), with the transcription activation properties of SMAD1 being extinguished by interaction with Klf4 (Morikawa et al., 2016). Instead, a major role of Bmp signalling appears to be through activation of Mek5, which activates Erk5 resulting in upregulation of the naïve factor Klf2 (Morikawa et al., 2016).

### 1.2.3. Alternative pluripotent identities

As described above, teratocarcinomas can arise from germ cells in mouse embryos (Stevens, 1967), suggesting that these cells are either already pluripotent or can readily convert to a pluripotent state. Subsequent work has shown that primordial germ cells (PGCs) themselves are not pluripotent, differentiating only to gametes, and cannot integrate into mouse blastocysts (Matsui, 1998). However, they express many pluripotency-associated factors (Kurimoto et al., 2008), and rapidly convert to a pluripotent identity when cultured in serum-based medium supplemented with FGF, SCF and LIF (Matsui et al., 1992, 1991; Resnick et al., 1992).

Meanwhile, there was significant work put into identifying conditions that could maintain long-term self-renewing human ESCs. Pluripotent stem cells (PSCs) were derived from non-human primates

using serum-based medium with or without human LIF, and required a feeder layer of MEFs for self-renewal (Thomson et al., 1996, 1995). Utilizing the experience and knowledge gained from working with these primate PSCs, human PSC lines were derived from frozen embryos shortly afterwards (Thomson et al., 1998). All of these primate cell lines were very distinct from mouse ESCs, however. On explanting embryos in culture, the ICM forms a distinctive domed structure, with cells maintaining close contact in a 3-dimentional arrangement. As mouse ESC lines are established, the colonies return to this domed morphology after each passage. The primate lines, however, all form a flat epithelium without such tight contact between individual cells (Thomson et al., 1998, 1996, 1995). Pluripotent human EGCs were established from PGCs, and showed a compacted, rounded morphology more reminiscent of mouse ESCs (Shamblott et al., 1998); however, poor derivation and differentiation efficiencies limited their use (Turnpenny et al., 2003).

There were also significant differences in signal requirements between mouse ESCs and primate PSCs. As described above, mouse ESCs are dependent on LIF and BMPs to maintain their identity, whereas primate PSCs are unresponsive to LIF (Dahéron et al., 2004; Humphrey et al., 2004) and differentiate to trophectoderm on BMP induction (Xu et al., 2002). Instead, human PSCs are reliant on the TGFβ-family ligands Activin A or Nodal and bFGF for their maintenance (Figure 1.3) (Amit et al., 2004; Beattie et al., 2005; James et al., 2005; Vallier et al., 2005; Wang et al., 2005; C. Xu et al., 2005; R.-H. Xu et al., 2005). Activin-Nodal signalling contributes to the proliferation of mouse ESCs but is not essential (Ogawa et al., 2007), while autocrine FGF signalling is important in their exit from pluripotency (Kunath et al., 2007) and blockade of downstream MEK/ERK signalling favours their self-renewal (Figure 1.3) (Burdon et al., 1999).

Initially, it was assumed that this was the result of species differences. The post-implantation mouse embryo is unusual in forming an egg-cylinder instead of the bilaminar disk seen in most mammalian embryos, so a level of signal divergence is not unexpected. However, the derivation of mouse pluripotent cell lines with properties highly reminiscent of primate PSCs from the post-implantation mouse epiblast suggested that this was not the case. Cells dependent on FGF and Activin/Nodal and forming flattened colonies were readily derived from post-implantation mouse embryos, and were termed EpiSCs (Brons et al., 2007; Tesar et al., 2007); these are also known as primed cells, as their expression of lineage markers implies that they are primed for differentiation. Like primate cells, these were unresponsive to LIF (Figure 1.3), and EpiSCs showed similar transcriptional responses to human

PSCs on signal inhibition (Tesar et al., 2007). Furthermore, it was demonstrated that cells of the pre-implantation mouse embryo could be captured *in vitro* as EpiSCs (Najm et al., 2011) as the epiblast cells differentiate until they reach a post-implantation like identity. This can be recapitulated in the efficient differentiation of mouse ESCs to EpiSCs by changing growth conditions (Guo et al., 2009). Together, this showed that it was possible that primate pre-implantation epiblast cells had converted to a post-implantation state, and this was the cause for some the differences between primate PSCs and mouse ESCs.

### 1.2.4. Current ESC culture conditions

Following research into the role of various signalling cues in the self-renewal of mouse ESCs described above, defined culture conditions for their maintenance in the absence of differentiation were developed. Relying on two pharmacological inhibitors – PD03, a MEK1/2 inhibitor, and Chiron, a GSK3 inhibitor – and LIF (Figure 3.1), this permits growth of ESCs in the absence of serum or feeders with little to no background differentiation (Ying et al., 2008). These conditions also allowed derivation of ESCs from genetic backgrounds that were previously considered non-permissive as they rarely, if ever, yielded self-renewing cell lines in conventional growth conditions (Nichols et al., 2009). Further, 2iL conditions allowed derivation of naïve rat ESCs for the first time (Buehr et al., 2008; Li et al., 2008).

As described above, autocrine FGF signalling through MEK/ERK plays a significant role in the spontaneous differentiation of ESCs, and blocking this pathway with PD03 increases the propensity of cells to self-renew (Figure 1.3). LIF positively regulates naïve gene expression through activation of Stat3. By inhibiting GSK3, Chiron prevents the degradation of β-catenin. β-catenin is not essential for the maintenance of the naïve identity (Andersson-Rolf et al., 2017; Raggioli et al., 2014; Wray et al., 2011). However, it does inhibit the repressive Tcf3-Groucho complex which otherwise antagonises self-renewal (Figure 1.3) (Pereira et al., 2006). Additionally, β-catenin is required for cell adhesion and genomic stability (Andersson-Rolf et al., 2017; Raggioli et al., 2014). In other contexts, GSK3 also regulates mTORC1/2, but in mouse ESCs it appears that GSK3 does not affect mTORC activity (Sanchez-Ripoll et al., 2013), while mTORC1 may play a role in regulating GSK3 activity downstream of Akt (Li et al., 2018).

Figure 1.3

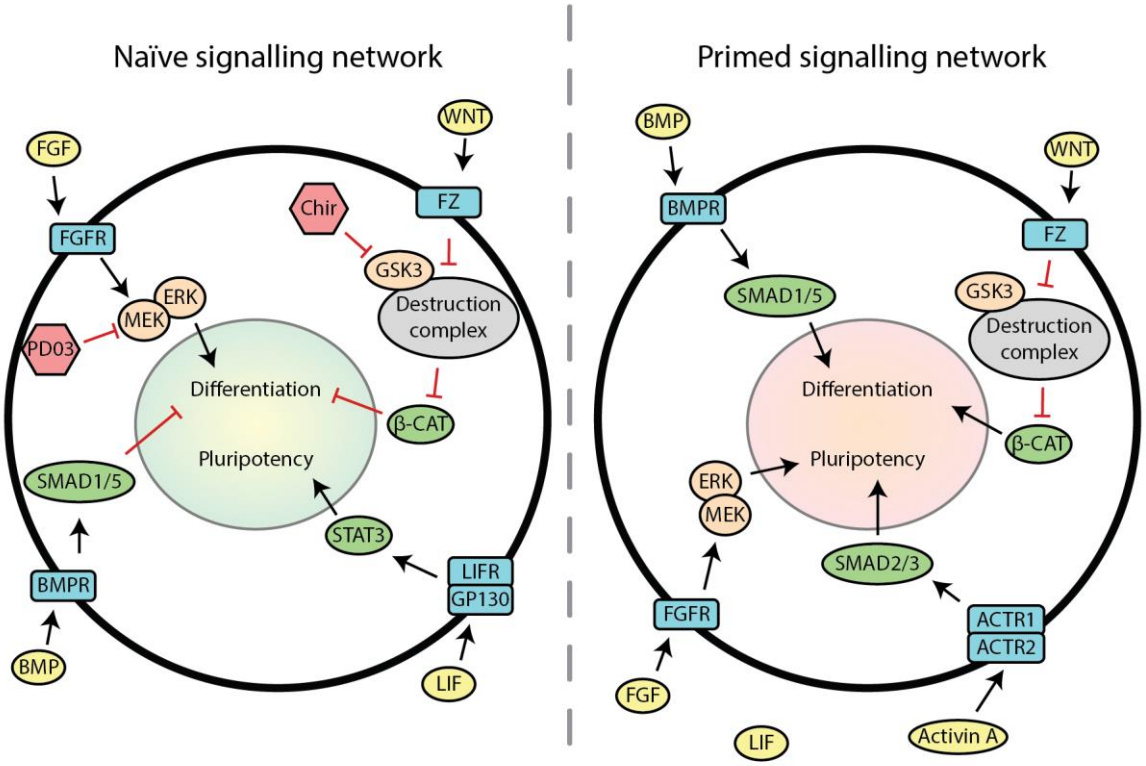



Figure 1.3 Summary of exogenous signals used to maintain naïve and primed pluripotency
Naïve and primed pluripotency rely on different signals to maintain self-renewal and induce differentiation. Naïve cells can be maintained in basal media containing BMP and LIF, or two components of 2iL (PD03, Chiron (Chir) and LIF). LIF induces activation of JAK/STAT signalling, leading to nuclear localisation of active STAT3, a coactivator of several pluripotency factors. BMP signalling causes activation of SMAD1 and 5, and induces binding of co-SMAD4; these upregulate inhibitors of differentiation. In parallel, BMP may upregulate Klf2 via MEK5/ERK5 signalling. WNT is not commonly used to promote naïve pluripotency. However, WNT signalling results in inhibition of the destruction complex that induces degradation of β-catenin. A similar effect is achieved by pharmacological inhibition of GSK3, a member of the complex, by Chiron. Downstream of this, β-catenin acts to inhibit TCF3, an inhibitor of naïve gene expression. FGF signalling via MEK/ERK induces differentiation of naïve cells; this can be prevented through pharmacological inhibition of MEK1/2 with PD03. Meanwhile, primed cells require MEK/ERK signalling to maintain pluripotency. TGF-β factors such as Activin and Nodal promote survival and self-renewal of primed cells via activation of SMAD2/3, which bind to co-SMAD4 to function as transcription factors. Contrary to their roles in naïve pluripotency, SMAD1/5 and β-catenin both promote the differentiation of primed cells downstream of BMP and WNT signalling respectively. Primed cells are largely unresponsive to LIF, having significantly reduced expression of LIFR. Note that these models are highly simplistic, as there is significant crosstalk between signalling pathways in reality.

In addition to providing an environment for robust self-renewal of ESCs, 2iL resolves an issue that was ever present in conventional ESC culture systems; cellular heterogeneity. In SL conditions, there is significant cell to cell variation in the expression of many pluripotency genes including Nanog (Chambers et al., 2007; Hatano et al., 2005), Esrrb (van den Berg et al., 2008) and Rex1 (Toyooka et al., 2008). A fraction of cells lose expression of Nanog due to the onset of differentiation, but others only transiently downregulate Nanog and then gradually return to a Nanog-high state (Chambers et al., 2007; Filipczyk et al., 2015). While in a Nanog-low state, however, cells have an increased

propensity to differentiate (Chambers et al., 2007; Kalmar et al., 2009), and may be biased in their differentiation (Canham et al., 2010; Singh et al., 2007). This is particularly troubling since expression of pluripotency genes appears to be relatively consistent between naïve epiblast cells in the mouse embryo (Guo et al., 2010). 2iL culture increases the level of Nanog and other pluripotency factors at a population level, partly by greatly reducing the amount of heterogeneity in their expression (Hastreiter et al., 2018; Kumar et al., 2014; Marks et al., 2012; Wray et al., 2011).

Mouse ESCs cultured in 2iL conditions also more closely mirror the epigenetic and transcriptional state of the naïve epiblast than cells grown in SL. *In vivo* these cells are characterized by global DNA hypomethylation (Howlett and Reik, 1991; Lane et al., 2003; Smith et al., 2012), almost entirely lacking 5-methylcytosine except at imprinted loci (Hanel and Wevrick, 2001; Reik and Walter, 2001). However, ESCs in SL show a significantly higher level of DNA methylation (Leitch et al., 2013). This is reverted to embryo-like levels of hypomethylation when cells are grown in 2iL medium (Leitch et al., 2013; Sim et al., 2017). More recently, RNA-sequencing has been performed on early mouse embryos and ESCs maintained in different conditions. Cells grown in 2iL closely resemble the preimplantation naïve epiblast, whereas SL cultured cells are far more heterogeneous and fail to reliably cluster with any single embryonic identity, more closely matching a post-implantation identity (Boroviak et al., 2015, 2014).

Interestingly, there have been recent reports of culture environments that support the growth of totipotent cells. However, it appears that these are most similar to ESCs with broader differentiation capacity induced by a more permissive epigenetic environment rather than recapitulating early embryonic identities *in vitro* (Yang et al., 2017).

More recently, conditions have been developed for the propagation of human PSCs in a state that much more closely resembles the pre-implantation embryo (Guo et al., 2016; Takashima et al., 2014; Theunissen et al., 2016, 2014). However, it remains unclear whether these conditions are optimal and how closely they recapitulate their *in vivo* counterparts (Bates and Silva, 2017). Interestingly, recent single-cell RNA sequencing data appears to show that a naïve compartment does exist within the pre-implantation embryo in rhesus monkey (D. Liu et al., 2018) and human (Stirparo et al., 2018), suggesting that these naïve-like human PSCs may be biologically relevant. Nonetheless, mouse ESCs still present a more attractive model system for studying naïve pluripotency factors and cell fate transitions for the time being.

### 1.2.5. Reprogramming to naïve pluripotency

In a landmark discovery, Takahashi and Yamanaka found that somatic cells could be reprogrammed back to a pluripotent identity in 2006. Overexpression of four factors – Oct4, Sox2, Klf4 and cMyc – was sufficient to convert fibroblasts to an induced pluripotent identity (Takahashi and Yamanaka, 2006), leading to the name iPSCs. While the original report did yield pluripotent cells, as indicated by the generation of tissues representing all three germ layers in teratomas, they were clearly distinct from ESCs; they continued to express retroviral transgenes to high levels, they only partially reverted to an ESC epigenome, and most lines lacked expression of key naïve markers such as Nanog and Rex1. The following year, it was reported that selection for the small fraction of cells that upregulated Nanog was sufficient to produce iPSCs with more ESC-like characteristics (Okita et al., 2007). Most of these cell lines were transgene independent as retroviruses were silenced on complete reprogramming. These naïve iPSCs were competent to generate chimeras and even contributed to the germ line. Since then, a large amount of effort has been put into developing protocols and identifying additional transgenes that contribute to more efficient derivation of iPSC lines, as well as finding alternative reprogramming cocktails and genes essential to the reprogramming process. Interestingly, the most irreplaceable factor appears to be Oct4, which is used itself or replaced by a factor that directly induces Oct4 in the majority of reprogramming cocktails (Radzisheuskaya and Silva, 2014).

Generally, somatic cell reprogramming is a highly inefficient process. One factor that contributes to this is the fact that the majority of cells stall at a stable, highly proliferative state prior to entering the naïve state. Application of 2iL greatly improves the efficiency of conversion of these pre-iPS intermediates to bona fide naïve iPSCs (Silva et al., 2008).

It is also possible to reprogram primed EpiSCs back to naïve pluripotency. This is a simpler process, likely owing to the more similar identities of the two cell types; both naïve iPSCs and primed EpiSCs express key pluripotency factors such as Nanog and Oct4. While it has been reported that EpiSCs can revert to a naïve identity of culture in 2iL conditions (Bao et al., 2009), this appears to be a peculiarity of a small number of cell lines. More often, it is reported that overexpression of one of many single naïve transcription factors is sufficient to achieve reprogramming of these cells (Gillich et al., 2012; Guo et al., 2009; Guo and Smith, 2010; Hanna et al., 2009; Silva et al., 2009; Zhang et al., 2016), however. Alternatively, activation of Stat3 signalling, which is typically inactive in EpiSCs, or chemical

alteration of the epigenetic landscape can mediate this identity change in mouse cells (Murayama et al., 2015; Ware et al., 2014; Yang et al., 2010; Zhang et al., 2016).

## 1.3. The naïve transcriptional network

### 1.3.1. Signalling and transcription regulation

The state of naïve pluripotency in mouse is maintained by a complex, highly redundant and interconnected transcriptional network. A plethora of transcription factors regulate the expression of one another to create a semi-stable network while repressing factors that would induce differentiation. As we have seen, in SL conditions secretion of FGF4 is a key role in the induction of differentiation. However, in the presence of MEK inhibitor, stimulation of LIF/STAT3 or WNT/β-catenin signalling is sufficient to maintain the self-renewal of undifferentiated ESCs. Knockout studies revealed a triad of factors that are required for the establishment of the naïve epiblast *in vivo*; Oct4, Sox2 and Nanog (Chambers et al., 2003; Mitsui et al., 2003).

Nanog is a homeodomain transcription factor. When overexpressed, ESCs are capable of self-renewal in the absence of LIF. However, Nanog is not the primary downstream effector of LIF/STAT3 signalling as demonstrated by the fact that hyperactivation of STAT3 does not affect Nanog expression, while constitutive Nanog expressing ESCs self-renew more effectively in the presence of LIF. It has been shown that Nanog acts synergistically with activated STAT3 to induce the pluripotency factor Klf4 (Hall et al., 2009; Li et al., 2005; Stuart et al., 2014). Nanog null ESCs can be maintained, albeit with significantly impaired self-renewal and a greater propensity for differentiation (Chambers et al., 2007). The transcription factor Esrrb appears to be downstream of Nanog. Constitutive expression of Esrrb rescues the proliferative and self-renewal defects of $Nanog^{-/-}$ ESCs and maintains self-renewal in the absence of LIF when overexpressed, mirroring the activity of Nanog (Festuccia et al., 2012). Like Nanog, $Esrrb^{-/-}$ cells can be maintained in the presence of LIF. Esrrb is a major effector downstream of GSK3 inhibition in mediating enhanced self-renewal. GSK3 forms part of the destruction complex that leads to degradation of β-catenin. A major function of β-catenin in ESCs is to inactivate the negative pluripotency regulator TCF3 by competing with the repressor Groucho for binding. TCF3 binds directly to Esrrb enhancer elements, Esrrb overexpression rescues differentiation induced by overexpression of TCF3, and Esrrb is permissive for self-renewal in the absence of the GSK3 inhibitor Chiron, while $Esrrb^{-/-}$ cells cannot maintain naïve pluripotency in media containing Chiron and only one of LIF or PD03 (Martello et al., 2012). Additionally Klf2,

Nanog, Nr0b1 and Tcfcp2l1 appear to be repressed by TCF3, and are thus responsive to treatment with Chiron (Martello et al., 2012). Tcfcp2l1 is also strongly responsive to LIF (Martello et al., 2013) as well as to PD03 (Ye et al., 2013), and overexpression of Tcfcp2l1 is sufficient to maintain naïve pluripotency in serum based media lacking any of the 2iL components (Ye et al., 2013). Notably, however, neither Oct4 nor Sox2 show strong, direct responses to any of these signals.

Both Oct4 and Sox2 are essential for the establishment of the naïve pluripotent epiblast *in vivo* (Avilion et al., 2003; Nichols et al., 1998) and for maintaining naïve ESCs *in vitro* (Avilion et al., 2003; Masui et al., 2007; Niwa et al., 2005, 2000). In the absence of either factor, mouse ESCs differentiate to a trophoblast-like identity (Masui et al., 2007; Niwa et al., 2000). Oct4 is a POU-class transcription factor, and binds to DNA via a homeodomain and a POU-specific domain, while Sox2 is an HMG-box transcription factor. They co-bind many genomic loci, interacting primarily in the presence of DNA (Fong et al., 2011; Kuroda et al., 2005; Lam et al., 2012; Merino et al., 2014). Notably, Sox2 binds to the minor groove, inducing a bending of the DNA which appears to play a functional role in the regulatory ability of Sox2 (Scaffidi and Bianchi, 2001). It has been suggested that Sox2 is important for maintaining expression of Oct4 via inducing positive regulators such as Nr5a2 and repressing negative regulators such as Nr2f2 (Masui et al., 2007). This is was reinforced by the discovery that constitutive expression of Oct4 can rescue $Sox2^{-/-}$ ESCs (Masui et al., 2007). The Oct4 enhancer is also bound by a large array of naïve transcription factors, implying highly redundant regulation (Chen et al., 2008; Young, 2011), which may explain why it is relatively insensitive to changes in expression of most individual pluripotency factors. Oct4 and Sox2 physically interact on DNA at many regulatory elements. Interestingly, some single-molecule imaging work has been performed examining the interactions of Oct4 and Sox2 with DNA both individually and together (Chen et al., 2014). This work found a limited effect of Oct4 on Sox2 DNA binding kinetics. However, in cells coexpressing Oct4 and Sox2, Oct4 spent a greater fraction of time bound at Oct4-specific binding sites versus non-specific binding relative to cells expressing Oct4 alone. Chen et al. interpret this as indicating ordered binding, with Sox2 recruiting Oct4 to their cooccupied loci. However, this view has recently been questioned (Biddle et al., 2019), with Biddle et al. reanalysing the data of Chen et al. and arriving at the conclusion that the interactions of Oct4 and Sox2 with DNA are more complex. While increased Sox2 does lead to increased binding of Oct4, increased Oct4 levels actually reduce the overall binding of Sox2 to DNA. They conclude that Oct4 and Sox2 may bind cooperatively at some loci and competitively at others, or

that energy expenditure may be required to explain the interaction of Oct4 and Sox2 with DNA; this might suggest, for example, that chromatin modifiers or nucleosome remodelling could play an important role in the way that Oct4 and Sox2 are recruited to the genome.

Co-binding of Oct4 and Sox2 is known to promote expression of some pluripotency factors such as Nanog and Klf4, but also negative regulators such as Fgf4 (Ambrosetti et al., 1997), while Oct4 and Nanog induce expression of Tcf3 (Yi et al., 2008). However, overexpression of either Oct4 or Sox2 promotes differentiation seemingly by destabilising the naïve network itself since their direct targets become rapidly downregulated (Ben-Shushan et al., 1998; Boer et al., 2007; Kopp et al., 2008; Niwa et al., 2000; Thorold W. Theunissen et al., 2011).

### 1.3.2. Previous models explaining the activity of Oct4 in naïve pluripotency

Despite years of study, it remains unclear exactly how Oct4 regulates and is required for naïve pluripotency. A large number of phenotypes have been documented following altered expression of Oct4 and its partners but as yet the progression from wild-type cells following perturbations, and how these disparate phenotypes relate to one another, remain relatively unexplored and poorly understood.

Loss of Oct4 is incompatible with establishing (Nichols et al., 1998) or maintaining (Niwa et al., 2000) naïve pluripotency, with cells exhibiting differentiation to a trophectoderm identity. Unlike factors such as Nanog and Esrrb, Oct4 does not contribute to a more robust pluripotency network on overexpression (Niwa et al., 2000). Instead cells are forced into differentiation. Furthermore, unlike these more conventional pluripotency factors, constitutive expression of Oct4 cannot prevent differentiation induced by removal of exogenous pluripotency signals such as LIF (Niwa et al., 2000). Therefore, it does not appear that Oct4 is simply acting downstream of cytokines to promote the maintenance of pluripotency, instead having a more complicated role.

It has previously been proposed that Oct4 may have an as-yet-unidentified partner, present at sufficiently low levels that in wild-type cells it is recruited to and activates a core set of pluripotency genes, but on overexpression of Oct4 it is spread too thinly across the genome as a result of Oct4 leading it to novel binding sites (Chambers, 2004; Niwa et al., 2002). As a result, expression of the core naïve factors is quenched and cells cannot maintain pluripotency. With this in mind, it is interesting to note that overexpression of Nanog is sufficient to rescue self-renewal in mouse ESCs overexpressing Oct4 to a modest level (Thorold W. Theunissen et al., 2011), a phenotype associated with this

hypothetical partner. Additionally, recent work has demonstrated enhanced self-renewal at the detriment of differentiation capacity in cells with reduced Oct4 levels (Karwacki-Neisius et al., 2013; Radzisheuskaya et al., 2013). This also matches predictions of such a model.

However, other models for the mechanism by which Oct4 maintains naïve pluripotency do exist. Notably, one model that successfully predicted the outcome of several reprogramming strategies proposes that pluripotency exists in a somewhat precarious state maintained by the balanced expression of a cohort of antagonistic lineage specifiers (Loh and Lim, 2011). It has been noted that overexpression of Oct4 causes mesendodermal differentiation (Niwa et al., 2000; Thomson et al., 2011). On the other hand, it has been shown that Sox2 acts to inhibit the formation of mesendoderm, while overexpression induces differentiation towards neural lineages (Thomson et al., 2011).

Evidence in favour of this model comes from reprogramming assays in which it has been shown that alternative mesendoderm specifiers or depletion of key ectodermal genes can replace exogenous Oct4 (Shu et al., 2013). Conversely, Sox2 can be replaced by other ectodermal specifiers. This culminated in the establishment of a novel cocktail of Gata6, Gmnn, Klf4 and cMyc which was capable of inducing somatic cell reprogramming (Shu et al., 2013). Of course, in the latter stages of reprogramming endogenous Oct4 and Sox2 are expressed, and since these are required in iPSCs it is impossible to test whether they are required for the final establishment of the naïve identity.

However, some of the phenotypes described above do appear to be evidence against such a model. For one, expression of alternative lineage specifiers has not been shown to rescue differentiation induced by knockout of Oct4 or Sox2. On the contrary, it has been demonstrated that Oct4 itself can rescue self-renewal on knockout of Sox2 (Masui et al., 2007), though it is possible that other lineage specifiers are present to counteract Oct4 in this scenario. Furthermore, it has been demonstrated that ESCs can readily differentiate to alternative identities following overexpression of Oct4, and that the culture conditions may be the main driving factor behind fate choice in these cells (Shimozaki et al., 2003; Simandi et al., 2016). Nonetheless, this model has not been entirely dismissed, and notably there is some preliminary evidence that natural fluctuations in the relative levels of Oct4 and Sox2 may impact cell fate decisions (Strebinger et al., 2018). Thus it is possible that the antagonistic role of pluripotency factors in lineage specification may contribute to the mechanism of pluripotency maintenance. Alternatively, it may be that this model closer matches the regulation of more advanced state of primed

pluripotency in which heterogeneous expression of lineage markers is commonly observed (Song et al., 2016; Tsakiridis et al., 2014).

It is interesting to observe that pluripotency appears to be regulated differently in other contexts. Prior to the establishment of the naïve epiblast, totipotent cells express many pluripotency factors such as Oct4 and Nanog but regulation of their expression differs. Knockout of either of these factors does not impair initial expression of the other in the early ICM (Le Bin et al., 2014), suggesting that their co-regulation is not yet established. However, cells lacking either factor fail to specify a naïve epiblast (Mitsui et al., 2003; Nichols et al., 1998). It may be that expression of these genes is more tightly linked to proliferation prior to the establishment of the blastocyst, though studies of the transcriptional network in blastomeres of the morula are lacking due to the absence of an *in vitro* model.

The regulation of pluripotency and the nature of the underlying transcription factor network is poorly described in post-implantation-like EpiSCs, but it is known that Oct4 expression is driven by a different enhancer element than in naïve cells (Choi et al., 2016; Tesar et al., 2007; Yeom et al., 1996). The functional consequences and differences in regulation as a result of this enhancer switch remain unclear, however. It appears that by E7.5, Sox2 is not required for expression of Oct4 (DeVeale et al., 2013). Additionally, Oct4 appears to have a different function in the post-implantation embryo than in naïve cells, promoting proliferation of the primitive streak and reducing p53 expression while promoting mesodermal over neural differentiation (DeVeale et al., 2013).

## 1.4. X chromosome inactivation

Perhaps the most striking example of the unique transcriptional and epigenetic landscape of naïve pluripotent cells is the presence of two active X chromosomes in females. In order to achieve equal expression of X-linked genes in males and females, most female cells inactivate one X chromosome. In cells of the early embryo, prior to establishment of the naïve epiblast, and later in the extraembryonic tissues the paternal X chromosome is always silenced (Figure 1.4). However, cells in the embryo proper display inactivation of either the maternal or paternal X, with the inactive chromosome being chosen individually in each cell around the time of implantation (Figure 1.4). Studies of XXX and XXXX individuals indicate that all but one X chromosome is inactivated, and similarly XXY males randomly inactivate one X chromosome (Grumbach et al., 1963). This demonstrates that cells possess a capacity to 'count' X chromosomes in order to determine whether random X chromosome inactivation (XCI) needs to occur. Analysis of tetraploid cells reveals that these

maintain two active and two inactive X chromosomes (Webb et al., 1992), leading the hypothesis that cells maintain a number of active X chromosomes equal to half their ploidy; two sets of autosomes are present for each active X chromosome.

The X inactivation centre (XIC) is a region of the X chromosome that must be present on multiple chromosomes in order for XCI to occur. Studies of cells with X chromosome truncations or deletions found that absence of this specific region led to a failure to inactivate an X chromosome in females (Rastan and Robertson, 1985). Interestingly, this region harbours an unusual gene, Xist, which is specifically expressed from the inactive X chromosome in differentiated cells (Borsani et al., 1991; Brockdorff et al., 1991). Xist is a long non-coding RNA (lncRNA) (Brockdorff et al., 1992) which is weakly expressed in ESCs but becomes strongly upregulated on their differentiation, specifically from the allele of the presumptive inactive X. As it becomes highly expressed, Xist RNA coats the entire X chromosome from which it is transcribed (Clemson et al., 1996). Deletion of Xist shows that it is absolutely required for silencing *in cis*, though $Xist^{+/-}$ female cells still inactivate the wild-type chromosome (Penny et al., 1996).

### 1.4.1. Imprinted XCI

It appears that regulation of non-random XCI is relatively straightforward. Initially on fertilization, all chromosomes are inactive until zygotic genome activation (Figure 1.4). However, the maternal Xist locus is imprinted and thus remains silent when transcription begins globally (Goto and Takagi, 2000; Kay et al., 1994). In males, Xist is therefore not expressed and the X chromosome remains active. Meanwhile in female cells, Xist is expressed from the paternal X chromosome resulting in silencing of the paternal X in all cells of the early embryo as described above (Figure 1.4). This non-random silencing persists throughout the tophectoderm and extraembryonic endoderm. However, the paternal X becomes reactivated in the naïve epiblast (Silva et al., 2009), resulting in cells with two active X chromosomes (Figure 1.4). At the same time, the imprints that silenced the maternal Xist locus are removed. This sets the stage for random XCI which occurs during implantation; the regulation of this process is far more complex and poorly understood.

Figure 1.4

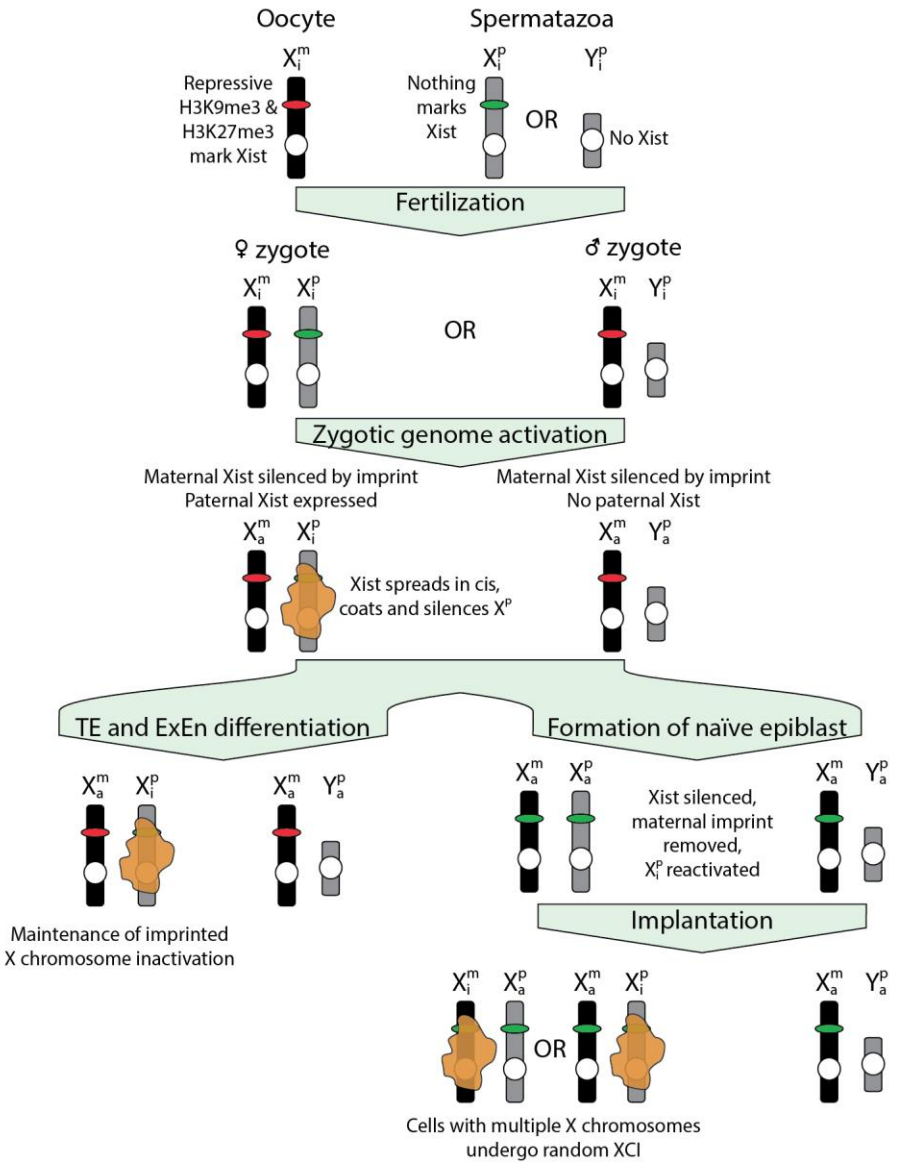



Figure 1.4 Outline of the processes of imprinted X inactivation, reactivation, and random XCI
Oocytes contain an X chromosome ($X^m$) with an imprinted Xist allele (red), while spermatozoa either carry a Y chromosome which lacks Xist, or an X chromosome ($X^p$) with no imprinting at the Xist locus (green). On fertilization, a male or female zygote is formed. Globally, there is no expression until the onset of zygotic genome activation. Xist is not expressed from $X^m$ due to the inherited epigenetic imprinting. Consequently, Xist is not expressed in male embryos. In female embryos, Xist is expressed from the Xist allele on the unmarked $X^p$. This coats the paternal X chromosome, leading to gene silencing and imprinted chromosome inactivation ($X^p_i$). Expression of Xist is silenced in cells of the established naïve epiblast; extraembryonic cells maintain imprinted inactivation of $X^p$. As a result, $X^p$ is reactivated in naïve cells, resulting in two active X chromosomes. Around the same time, the imprints at the Xist locus on $X^m$ are removed. During implantation and exit from the naïve state, cells count the number of X chromosomes, and upregulate Xist from a random allele if more than one X chromosome is present. Again, this coats the chromosome *in cis*, resulting in random XCI. Note, female naïve ESCs and iPSCs exhibit two active X chromosomes like cells of the naïve epiblast and recapitulate the events of random XCI on differentiation.

## 1.4.2. Elements involved in counting, choice, and initiation of XCI

Integration of multiple copies of a 450kb region, including the Xist locus, into an autosome in male ESCs was sufficient to induce early events of XCI upon their differentiation, with expression of Xist

being induced from both the endogenous and transgenic loci in a large fraction of cells (Lee et al., 1996). It is worth noting that differentiating female cells only express Xist from a single locus, that of the chromosome undergoing XCI; however, there are multiple tandem integrations of the XIC in this system. By analogy to individuals with trisomy or tetrasomy X, the multiple XIC integrations appear to be counted separately leading to expression from several transgenic Xist loci as well as the endogenous locus. This demonstrates that the XIC contains elements required for counting. Remarkably, multiple autosomal integrations of a far smaller transgene consisting of the Xist locus and a total of 15kb of flanking sequences was also capable of inducing transgenic and endogenous Xist expression (Herzing et al., 1997). Additionally, these cells demonstrated local silencing of markers within the transgene demonstrating that Xist can induce gene inactivation *in cis*.

Once counting has occurred and it has been determined that XCI should occur, a choice must be made such that one, and only one, X chromosome is inactivated in female cells. Offspring of mice from different genetic backgrounds show a bias in the choice of X to silence. This is conventionally defined as the inheritance of different X-linked alleles at the X-chromosome controlling element (Xce). Different alleles have different strengths such that a chromosome carrying the $Xce^{a}$ allele is more likely to be inactivated than one carrying $Xce^{b}$, while a chromosome carrying $Xce^{d}$ is the least likely to be inactivated. Phenotypically, this results in up to 75% of cells inactivating the X chromosome harbouring the weakest ($Xce^{a}$) allele in $Xce^{a/c}$ heterozygous mice (Fowlis et al., 1991). Genetic mapping of the Xce locus has not revealed a single polymorphic gene, but instead a broad, repetitive region several hundred kb downstream of the Xist locus (Calaway et al., 2013; Simmler et al., 1993). It has been suggested that the cause of bias between different Xce alleles is the result of copy number variance for a trans-acting factor that regulates the initiation of XCI (Calaway et al., 2013), though this has not been demonstrated.

Three lncRNAs are located within the XIC on the opposite strand to Xist. Xite and Tsx are located ~20kb and ~36kb downstream of Xist respectively, whereas the antisense Tsix transcript overlaps Xist and extends beyond the gene in both directions. Xite and Tsx both positively regulate Tsix. Knocking out Xite reduces Tsix expression *in cis*; Expression of Xite does not appear to be required for this effect, so it may act as a conventional enhancer element (Ogawa and Lee, 2003). Similarly, knockout of Tsx results in reduced expression of Tsix in undifferentiated and differentiating ESCs (Anguera et al., 2011). Deletion of a 65kb region distal to exon 6 of Xist effectively knocks out Tsx, Xite and Tsix.

Differentiating female cells carrying this deletion on one X chromosome still results in upregulation of Xist and XCI; however, the chromosome harbouring the deletion is always chosen for inactivation (Clerc and Avner, 1998). However, after reintroducing Tsix and the 3' end of Xist to these cells, they still do not undergo random inactivation upon differentiation (Morey et al., 2001); they still always inactivate the chromosome carrying a now smaller deletion, suggesting that elements involved in choice are contained within this deletion region. That said, a small targeted deletion of ~4kb at the Tsix transcriptional start site is sufficient to induce non-random X inactivation (Lee and Lu, 1999), demonstrating that both Tsix and other factors in the region have this effect. Integration of a YAC carrying an extensive region including the XIC, but with a ~3kb deletion at the Tsix transcription start site, into male ESCs also induces non-random inactivation on differentiation (Debrand et al., 1999), with the autosome carrying the mutant XIC always being coated by Xist. Additionally, cells carrying the full 65kb deletion have a tendency to lose a significant portion of the wild-type X chromosome, resulting in a truncated chromosome lacking the XIC. Despite being effectively XO, these cells still undergo XCI implying that the counting mechanism is also broken (Morey et al., 2001). The same precocious X inactivation is observed in male cells carrying the 65kb deletion (Morey et al., 2004). This is not the case in male cells with the small Tsix inactivating deletion; as in wild-type cells, these silence Xist on differentiation (Lee and Lu, 1999). However, reintroducing a 37kb section, restoring the 3' end of Xist, the Tsix start site, and Xite, into cells carrying the 65kb deletion restores normal counting (Morey et al., 2004). Altogether, the regulation of this region is unclear, and the precise role of each component has yet to be reliably described.

Oct4, Nanog and Sox2 all bind within intron 1 of Xist (Navarro et al., 2008). Knockout of Nanog in male ESCs results in mild upregulation of Xist, though not to levels seen during X inactivation (Navarro et al., 2008). However, knockout of Oct4 results in upregulation of Xist to levels seen in differentiating female ESCs (Navarro et al., 2008). Despite this, removal of Xist intron 1 has little to no impact on the ability of cells to properly undergo XCI (Minkovsky et al., 2013), and siRNA knockdown of Oct4 appears to have only a mild effect of Xist expression (Donohoe et al., 2009; Minkovsky et al., 2013). There is some evidence that Oct4 and Sox2 bind within Tsix regulatory elements (Donohoe et al., 2009). Knockdown of Oct4 reduces expression of Tsix, while Sox2 knockdown increases expression (Donohoe et al., 2009). It appears that loss of Oct4 prevents pairing of X chromosomes in female cells, allowing a small number of cells to initiate XCI from both

chromosomes (Donohoe et al., 2009). However, the manner in which Oct4 achieves this, and exactly how binding of Oct4 affects Xist expression is unclear. Interestingly, X-reactivation also occurs in PGCs which, as previously discussed, possess a latent capacity for pluripotency and express many naïve-associated genes. However, it is unclear how this is regulated.

## 1.5. Aims of this work

Oct4 is present in embryonic development from the oocyte through to gastrulation. The level of Oct4 is tightly controlled within the embryo, and fluctuates very little prior to the exit of cells from pluripotency (Joo et al., 2014; Rosner et al., 1990; Torres-Padilla and Chambers, 2014). It is unclear why this specific level has been adopted; evidence shows that a reduced level of Oct4 is optimal for the self-renewal of naïve pluripotent cells *in vitro* (Karwacki-Neisius et al., 2013; Radzisheuskaya et al., 2013). Oct4 may act as regulator to the naïve state, ensuring that cells are capable of exiting pluripotency, as they must do so *in vivo* in order to progress in development. However, this is at odds with reports that Oct4 upregulates expression of naïve pluripotency genes (Ben-Shushan et al., 1998; Chew et al., 2005; Rodda et al., 2005). In this thesis I hope to address the impact of loss of Oct4 from naïve cells and of maintenance of Oct4 in somatic cells. In doing so, I hope to demonstrate that more than a naïve factor, Oct4 is a fate transition factor.

It is well established that Oct4 behaves in a manner highly dependent on the level of expression (Karwacki-Neisius et al., 2013; Niwa et al., 2000; Radzisheuskaya et al., 2013). In chapters 3 and 4, I seek to control Oct4 levels more tightly than has been performed in past work. With this in mind, I will investigate the impact of sustaining an ESC level of Oct4 on differentiation *in vivo* and on somatic cell reprogramming (chapter 3). Additionally, starting from a naïve identity, I will investigate the consequences of losing Oct4 without passing through a protracted Oct4-low state (chapter 4). In chapter 5, I will compare the differentiation induced by loss of Oct4 to alternative methods extraembryonic differentiation, particularly with regards to the (mis)regulation of Xist.

Conventionally, reprogramming to naïve pluripotency is achieved with very high levels of Oct4, potentially leading to neomorphic effects. Additionally, Oct4 is typically introduced via retroviral transfection, resulting in variable expression between cells. In this work, I will investigate the capacity of Oct4 to reprogram cells at a defined level, that of mouse ESCs.

Oct4 knockout experiments have been performed in the past, and result in differentiation of ESCs to a trophoblast-like identity (Niwa et al., 2000). However, it has recently been shown that cells expressing a low level of Oct4 become trapped in the naïve state (Karwacki-Neisius et al., 2013; Radzisheuskaya et al., 2013), and even overexpress pluripotency factors in some cases. Previous experiments have relied upon gradual degradation of Oct4 protein following ablation of the Oct4 gene or transcriptional repression (Nichols et al., 1998; Niwa et al., 2000). As a result, these cells will have transiently passed through an Oct4-low state, and it is unclear what effect this has on their progression. I will develop and utilize a rapid inducible system to achieve degradation of Oct4 at the protein level and investigate the order of events as these cells exit naïve pluripotency.

Additionally, I will use this system to investigate the impact of loss of Oct4 on the genomic binding of other transcription factors prior to transcriptional changes; rapid depletion of Oct4 makes it possible examine changes in transcription factor binding prior to identity changes. As a result, it is possible to directly ascribe changes to the removal of Oct4 using this novel system, whereas following gradual depletion of Oct4 differences in transcription factor localization could be caused by the onset of differentiation rather than due to loss of Oct4 specifically.

I will also investigate the impact of loss of Oct4 on Xist expression. First, following the observation that exit from pluripotency induces Xist expression (Sousa et al., 2018), I will examine the effect of differentiation of ESCs by depletion of Oct4 on the initiation of XCI. I will then investigate whether Xist expression is misregulated during extraembryonic differentiation induced by alternative means (chapter 5), and discuss the implications of this for the current models of XCI.

Together, this body of work primarily focuses on the activity of Oct4; the mechanism by which this factor contributes to the naïve pluripotent identity is unclear, and the contradictory phenotypes associated with different expression levels of Oct4 have led to significant confusion. I hope to draw together several strands of work, both current and historical, to establish a coherent model for the role of Oct4 within the naïve state. More generally, I will investigate the impact of Oct4 on cell fate transitions, as this protein has a much broader role than simply maintaining a pluripotent identity.

# Chapter 2: Materials and Methods

## 2.1. Cell Culture

### 2.1.1. Culture conditions

All cell types were manipulated in a sterile BioMAT Class II Microbiological Safety Cabinet (Thermo Scientific) and maintained in a humidified Sanyo incubator (MCO-18M) at 37°C and 7% $CO_2$, unless specified otherwise.

ESCs were cultured in N2B27 supplemented with CHIR99021 (Chiron), PD0325901 (PD03) and LIF (N2B27-2iL) or serum-based medium supplemented with LIF (SL) as specified, on tissue culture plastic (Falcon) coated with 0.15% gelatin (Sigma Aldrich) in DPBS (Sigma Aldrich). iPSCs were cultured in N2B27-2iL, SL, or GMEM supplemented with Knockout serum replacement (KSR), Chiron, PD03 and LIF (KSR-2iL) as specified, on gelatin-coated tissue culture plastic. NSCs were cultured in NSC medium on tissue culture plastic coated with 10 μg/ml fibronectin (Millipore) in DPBS. MEFs were cultured in SL on gelatin-coated tissue culture plastic. PLAT-E cells were cultured in SL on tissue culture plastic with no additional coating.

N2B27 was made as follows:
1 : 1 Neurobasal (Gibco) : DMEM/F12 (Gibco), 0.5% N2 (homemade: Tissue Culture facility, W-MRC CSCI, University of Cambridge), 1% B27 (Gibco), 2 mM l-glutamine (Gibco), 0.1 mM 2-mercaptoethanol (Gibco).

N2B27-2iL was made by supplementing N2B27 with 3 μM CHIR99021 (Stewart lab, Dresden), 1 μM PD0325901 (Stewart lab, Dresden) and 20 ng/ml mLIF (homemade: Department of Biochemistry, University of Cambridge).

TSC medium was made by supplementing N2B27 with 12.5 ng/ml FGF2 (homemade: Department of Biochemistry, University of Cambridge), 20 ng/ml Activin A (homemade: Department of Biochemistry, University of Cambridge), 10 μM XAV939 (Tocris), 5 μM Y27632 (Tocris).

SL was made as follows:

GMEM without l-glutamine (Sigma Aldrich), 10% FCS (Labtech), 2mM l-glutamine, 0.1mM 2-mercaptoethanol, 1X MEM non-essential amino acids (Sigma Aldrich), 1mM Sodium Pyruvate (Sigma Aldrich), 20 ng/ml mLIF.

KSR-2iL was made as follows:

GMEM without l-glutamine, 10% KOSR (Gibco), 1% FCS (Labtech), 2mM l-glutamine, 0.1mM 2-mercaptoethanol, 1X MEM non-essential amino acids, 1mM Sodium Pyruvate, 20 ng/ml mLIF, 3 μM CHIR99021, 1 μM PD0325901.

NSC medium was made as follows:

DMEM/F12, 0.5% N2, 1% B27, 29mM glucose (Sigma Aldrich), 4.5mM HEPES (Gibco), 1X MEM non-essential amino acids, 120 μg/ml BSA (Gibco), 0.1 mM 2-mercaptoethanol, 10 ng/ml Egf (Peprotech), 20 ng/ml FGF2.

Media were supplemented with 1X penicillin-streptomycin (Sigma Aldrich), 500 nM 4-hydroxytamoxifen (Sigma Aldrich), 1 μg/ml doxycycline (MP Biomedicals), 500 μM IAA (Cayman Chemical), 100 μg/ml cycloheximide (Sigma Aldrich), 0.5 μM 5-azacytidine (Sigma Aldrich), 400 μg/ml G418 (Life Technologies), 200 μg/ml hygromycin B (Life Technologies), 20 μg/ml blasticidin (Life Technologies), 1μg/ml puromycin (Sigma Aldrich) as required.

### 2.1.2. Passaging and freezing

ESCs were passaged every 2-4 days according to density and colony size. Briefly they were dissociated with pre-warmed TrypLE Express (Gibco), diluted 1:10 in DMEM/F12, pelleted at 300g for 3 minutes, supernatant was removed and cells were resuspended in growth medium and plated. Typically cells were split 1:6 each passage.

NSCs were passaged every 3-5 days according to density. Briefly, they were dissociated with pre-warmed accutase (Millipore), diluted 1:10 in DMEM/F12, pelleted at 300g for 3 minutes, supernatent was removed and cells were resuspended in growth medium and plated. Typically cells were split 1:4 each passage.

iPSCs were passaged every 2-4 days according to density and colony size. iPSCs were passaged as NSCs. Typically cells were split 1:6 each passage.

MEFs were passaged every 4-8 days according to density. MEFs were passaged as ESCs. Typically cells were split 1:6 each passage.

All cells were frozen in 9 : 9 : 2 DMEM/F12 : FCS : DMSO in a -80°C freezer before transfer to a liquid nitrogen tank for long-term storage.

### 2.1.3. Cell lines

$Oct4^{F/-}$ Rosa26:CreERt2 ESCs (from male and female littermates) and NSCs (male) were previously derived from embryos. Briefly, $Oct4^{+/-}$ mice were bred with $Oct4^{F/F}$ Rosa26:CreERt2 mice. NSCs were treated with 4-OHT to give $Oct4^{-/-}$ NSCs. These were transduced with retroviral Klf4 and cMyc, and transfected with PB-CAG-Oct4 to generate $Oct4^{-/-}$ fixed-Oct4 iPSCs (Radzisheuskaya et al., 2013). $Oct4^{-/-}$ NSCs were alternately transfected with pPB-CAG-NLS-m.OsTir1-IRES-bsd and either pPB-CAG-Oct4-PGK-hyg or pPB-CAG-O4AID-PGK-hph, and transduced with retroviral Oct4, Klf4 and cMyc to generate $Oct4^{-/-}$ Oct4 or Oct4-AID iPSCs (See 2.1.5 NSC Reprogramming).

$Oct4^{F/-}$ ESCs (Line A and Line B) were a generous gift from Aliaksandra Radzisheuskaya. They were transfected with linearized pCAG-$CreER^{T2-NLS}$-IRES-bsd

Inducible Oct4-AID ESCs were generated by transfecting wild-type E14Tg2a cells (a generous gift from Aliaksandra Radzisheuskaya) with pPB-CAG-rtTA-IRES-pac, pPB-CAG-NLS-m.OsTir1-IRES-bsd and pPB-TRE-O4AID-PGK-hph.

iOct4 cells were a generous gift from Aliaksandra Radzisheuskaya. Briefly, wild-type E14Tg2a cells were transfected with PB-TRE-Oct4 and PB-CAG-rtTA. After selection, a clonal line was derived.

$Oct4^{-/-}$ TetOFF-Oct4 ESCs were a generous gift from the Smith lab (ZHBTc4 cells, (Niwa et al., 2000)). $Oct4^{-/-}$ Oct4-AID ESCs were generated by transfecting these cells with pPB-CAG-NLS-m.OsTir1-IRES-bsd and pPB-CAG-O4AID-PGK-hph and maintaining cells in the presence of dox.

Rex1-GFP ESCs were a generous gift from the Smith lab (Wray et al., 2010). iCdx2 and iGata6 cells were generated by transfecting this cells with pPB-CAG-rtTA-IRES-pac and either pPB-TRE-Cdx2-PGK-hyg or pPB-TRE-Gata6-PGK-hyg.

### 2.1.4. Cell transfection

NSCs were transfected using Amaxa Nucleofection Technology (Lonza AG). 2 x $10^6$ cells were used per nucleofection, using program T-020. ESCs and iPSCs were transfected using Lipofectamine 2000

(Life Technologies). In this case, 350000 cells were plated the day prior to transfection, and on the day of transfection a mixture of Lipofectamine 2000, plasmid DNA and DMEM/F12 was added to fresh culture medium for 24 hours. Typically, piggyBac transposon (pPB) plasmids were co-transfected with piggyBac transposase expression vector pBase in a 1:1 ratio to generate stable cell lines. pCAG-CreER$^{T2\text{-}NLS}$-IRES-bsd was linearized using ScaI prior to transfection. PLAT-E cells were transfected using FuGENE6 (Promega).

### 2.1.5. NSC Reprogramming

Retroviruses were produced in PLAT-E cells; briefly, cells were transfected with pMXs-Oct3/4, pMXs-Klf4 or pMXs-cMyc using FuGENE 6 reagent (Promega). Medium was changed the following day, and 48 hours virus-containing supernatants were collected. The media were filtered through 0.22 μm filters and mixed in equal ratio, then polybrene (Sigma-Aldrich) was added to a final concentration of 4 μg/ml. The polybrene/virus mixture was applied to NSCs. 24 hours later NSCs were nucleofected with a 1:5 ratio of pBase and either pPB-CAG-O4AID-PGK-hph or pPB-CAG-Oct4-PGK-hyg using Amaxa Nucleofection Technology and plated in NSC medium for 2 days then switched to SL medium. Medium was switched to KSR-2iL, and selection was added for expression of the endogenous Oct4 locus on the ninth day in KSR-2iL. Once colonies had expanded they were passaged into N2B27-2iL.

### 2.1.6. Reprogramming somatic cells from E8.5 embryos

Chimeras were generated by blastocyst injection of fixed-Oct4 iPSCs by William Mansfield using standard methodology and C57BL/6 host embryos. All animal work was performed in accordance with Home Office guidelines and regulations at the University of Cambridge, UK. The posterior portion of E8.5 embryos was dissected away and discarded, and the anterior portion was manually dissociated in a small volume of PBS and cultured with or without addition of 5-aza in N2B27 supplemented with LIF or KSR medium with LIF but lacking 2i in wells coated in gelatin. After 7 days, 5-aza was removed where present and 2i was applied to all conditions. After 5 days G418 was added.

### 2.1.7. NSC and MEF derivation

Chimeras were generated as above. Embryos were harvested at E12.5 and visceral organs were removed and genital ridges were identified and carefully dissected away. The brain was dissected out from the skull and manually dissociated in a small volume of PBS. Material was then plated in NSC medium on laminin-coated wells. The remaining carcass was dissociated in prewarmed trypsin-EDTA

(Life Technologies) and manually dissociated. Trypsin was inactivated by addition of FCS and material was plated in SL medium on gelatin-coated wells.

## 2.2. Molecular Biology

### 2.2.1. Plasmids and cloning

Table 2.1 List of plasmids used in this thesis

| Plasmid name | Source |
|---|---|
| pCAG-CreER$^{T2-NLS}$-IRES-bsd | Kindly provided by Dr Joerg Betschinger |
| pMXs-Oct3/4 | Addgene plasmid #13366 |
| pMXs-Klf4 | Addgene plasmid #13370 |
| pMXs-cMyc | Addgene plasmid #13375 |
| pBase (CMV-PBase) | Kindly provided by Aliaksandra Radzisheuskaya |
| pPB-CAG-rtTA-IRES-pac | Kindly provided by Keisuke Kaji |
| pPB-CAG-rtTA-IRES-bsd | Kindly provided by Tim Lohoff |
| pCAGGS-NLS-m.OsTir1-IRES-pac | Kindly provided by Dr Oliver Baker |
| pPB-CAG-Dest-PGK-hph | Kindly provided by Dr Joerg Betschinger |
| pPB-TRE-Dest-PGK-hph | Kindly provided by Aliaksandra Radzisheuskaya |
| pDONR211 | Obtained from Life Technologies |
| * pEntr Oct4-AID | Synthesized by Life Technologies |
| * pEntr NLS-m.OsTir1 | |
| * pEntr Cdx2 | Synthesized by Life Technologies |
| * pEntr Gata6 | Synthesized by Life Technologies |
| * pPB-CAG-O4AID-PGK-hph | |
| * pPB-TRE-O4AID-PGK-hph | |
| * pPB-CAG-NLS-m.OsTir1-IRES-bsd | |
| pPB-CAG-Oct4-PGK-hyg | Kindly provided by Aliaksandra Radzisheuskaya |
| pPB-TRE-Oct4-PGK-hyg | Kindly provided by Aliaksandra Radzisheuskaya |
| * pPB-TRE-Cdx2-PGK-hyg | |
| * pPB-TRE-Gata6-PGK-hyg | |

* pEntr Oct4-AID, pEntr Cdx2 and pEntr Gata6 were synthesized and cloned into pDONR221 by Life Technologies. pEntr Oct4-AID was designed by taking the mouse Oct4 coding sequence from the Ensembl public database and adding the *Arabidopsis thaliana* Iaa17 coding sequence to the C-terminus separated by a short PG linker. pEntr Cdx2 and pEntr Gata6 use the relevant mouse coding sequences from the Ensembl public database. pEntr NLS-m.OsTir1 was generated by PCR amplifying the mouse-optimised, SV40-NLS tagged *Oryza sativa* Tir1 coding sequence from pCAGGS-NLS-m.OsTir1-IRES-puro using primers to introduce attB Gateway cloning arms; the PCR product was incubated with pDONR211 and BP Clonase II enzyme mix (Life Technologies) according to the

manufacturer's instructions and transformed into cloning grade chemically competent E.cloni bacteria (Lucigen). pPB-CAG-GOI and pPB-TRE-GOI expression vectors were generated incubating the relevant pEntr and Dest vectors with LR Clonase II enzyme mix (Life Technologies) according to the manufacturer's instructions; after proteinase K was added to stop the reaction, 2μl of the reaction mix was transformed into cloning grade chemically competent E.cloni bacteria.

### 2.2.2. RT-qPCR and High Throughput RNA Sequencing

Total RNA was isolated from cultured cells using the RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. Cells were dissociated then pelleted at 300 g for 3 mins and supernatant was removed. Cell pellets were stored at -80 °C until they were processed further. RNA preparation included homogenisation of cell lysate with QIAshredder columns (Qiagen) and on-column DNA digestion with RNase-free DNase I (Qiagen). RNA quantity and purity were assessed using a Nanodrop ND-1000 spectrophotometer. For RT-qPCR, RNA was reverse-transcribed using SuperScript III First-Strand Synthesis SuperMix for RT-qPCR (Life Technologies); where possible 1 μg RNA was used for all samples, otherwise equal quantities of RNA was used for all samples of an experiment. For high throughput RNA sequencing, RNA integrity was assessed on a Qubit Fluorometer (ThermoFisher Scientific) and Agilent Bioanalyzer Nano Chips (Agilent Technologies). Depletion of ribosomal RNA was performed on 2-5μg of total RNA using the Ribo-Zero rRNA Removal Kit (Illumina) and libraries were produced from 10-100ng of ribosomal-depleted RNA using NextFlex Rapid Directional RNA-seq Kit (Bioo Scientific) with 12 cycles of PCR amplification. Libraries were pooled in equimolar quantities and sequenced on the HiSeq4000 platform (Illumina) at CRUK. Library preparation was performed by the W-MRC CSCI genomics facility.

### 2.2.3. List of primers and Taqman probes

Table 2.2 Pre-designed taqman probes

| Target | ID |
|---|---|
| Gapdh (VIC) | 4352339E |
| Cdx2 | Mm01212280_m1 |
| Esrrb | Mm00442411_m1 |
| Gata4 | Mm00484689_m1 |
| Klf4 | Mm00516104_m1 |
| Lifr | Mm00442942_m1 |
| Nanog | Mm02384862_g1 |
| Nr0b1 | Mm00431729_m1 |

| | |
|---|---|
| Oct4 | Mm00658129_gH |
| Pax6 | Mm00443081_m1 |
| Rex1 | Mm03053975_g1 |
| Sox2 | Mm03053810_s1 |
| Xist | Mm01232884_m1 |

Table 2.3 Custom Taqman probes

| Target | Primer | Sequence |
|---|---|---|
| Retroviral cMyc | Fw | TGGTACGGGAAATCACAAGTTTGTA |
| | Rv | GGTCATAGTTCCTGTTGGTGAAGTT |
| | Probe | CCCTTCACCATGCCCC |
| Retroviral Klf4 | Fw | TGGTACGGGAAATCACAAGTTTGTA |
| | Rv | GAGCAGAGCGTCGCTGA |
| | Probe | CCCCTTCACCATGGCTG |
| Retroviral Oct4 | Fw | TGGTACGGGAAATCACAAGTTTGTA |
| | Rv | GGTGAGAAGGCGAAGTCTGAAG |
| | Probe | CACCTTCCCCATGGCTG |

Table 2.4 RT-qPCR Sybr primers

| Target | Primer | Sequence |
|---|---|---|
| Elf5 | Fw | CCCTCCTCCTCTTCAAAACC |
| | Rv | AAGTTGCCACAAGACCATCC |
| Gata6 | Fw | CCCACTTCTGTGTTCCCAATTG |
| | Rv | TTGGTCACGTGGTACAGGCG |
| Gapdh | Fw | CCCACTAACATCAAATGGGG |
| | Rv | CCTTCCACAATGCCAAAGTT |

Table 2.5 ChIP qPCR primers

| Target | Primer | Sequence |
|---|---|---|
| Oct4 distal enhancer | Fw | GCATAACAAAGGTGCATGATAGCT |
| | Rv | AAATAAAGGCAGCGACTTGGAA |
| Sox2 proximal | Fw | CAGGTTCCCCTCTAATTAATGCA |
| downstream enhancer | Rv | CATTACCACGTGAATAATCCTATATGC |
| Nanog enhancer | Fw | CCCACCTGTCCCTAGTCCCCGCT |
| (OSN peak) | Rv | TTGGAACTAGCTGTGTGGGTGGGG |
| Chr6 negative locus | Fw | ACTACCCAAACTATTGCTCCTGA |
| | Rv | GCTTAACCTGCTCTCCCAGG |
| Klf4 enhancer | Fw | TGTCCTCTCCACTCCCACAA |
| | Rv | AGGAGTGACTGCGTCAAACA |
| Nanog enhancer | Fw | CACCTCTTCGCTCGGATCTT |

| | | |
|---|---|---|
| (H3K27ac peak) | Rv | CTCCGGGTCAAAGGAGTCTG |
| Rex1 enhancer | Fw | CTCAATGTGCATCTACTGTGCTC |
| | Rv | ACTGCATCGATCACCAGCTAC |

### 2.2.4. Chromatin Immunoprecipitation (ChIP)

ChIP was performed as followed:

Cells ($5\times10^6$ per IP for histone modifications or $10\times10^6$ for transcription factors) were fixed for 10 mins in Fixing Solution (5 mM HEPES pH 7.5, 10 mM NaCl, 0.1 mM EDTA, 50 μM EGTA, 0.4% formaldehyde for histone modifications or 1% formaldehyde for transcription factors, with protease inhibitors). Fixation was halted by addition of an excess of glycine. Cells were pelleted at 1400 rcf and washed with 1X ice cold PBS. Pellets were frozen at this point and stored at -80 °C.

Nuclei were isolated: Cells were incubation with Lysis Buffer 1 (50 mM HEPES pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% Glycerol, 0.5% NP40, 0.25% Tx100) at 4 °C for 10 mins, pelleted, then incubated with Lysis Buffer 2 (10 mM Tris pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA) at 4 °C for 10 mins.

Nuclei were pelleted then resuspended in Shearing Buffer (50 mM Tris pH 8.0, 1% SDS, 10 mM EDTA) and sonicated to obtain an average DNA fragment size of 300-600 bp depending on the protein of interest. Debris was removed by centrifugation at 14000 rcf at 4 °C, and a sample was taken to determine fragment size. Sonicated chromatin was diluted 1:10 in Dilution Buffer (50 mM Tris pH 8.0, 167 mM NaCl, 1.1% Tx100, 0.11% Sodium deoxycholate). Dynabeads Protein G (Invitrogen) were prepared by pre-incubating with isotypic IgG antibody. Diluted chromatin was pre-cleared by incubating with the beads for 2 hr at 4 °C. Supernatent was collected, a 10% input sample was taken for relative quantitation, and the chromatin was the incubated overnight at 4 °C with 1-5 μg of antibody or 2 μg of isotypic IgG (see Table 2.6). Beads were blocked overnight with BSA. The chromatin-antibody mix was incubated for 1 hr with BSA-blocked beads at 4 °C to allow chromatin-conjugated antibody to bind to the beads. Beads were then washed twice in Wash Buffer 1 (50 mM Tris pH 8.0, 0.1% SDS, 0.1% Sodium deoxycholate, 1% Tx100, 150 mM NaCl, 1 mM EDTA, 0.5 mM EGTA), once in Wash Buffer 2 (50 mM Tris pH 8.0, 0.1% SDS, 0.1% Sodium deoxycholate, 1% Tx100, 500 mM NaCl, 1 mM EDTA, 0.5 mM EGTA), once in Wash Buffer 3 (50 mM Tris pH 8.0, 250 mM LiCl, 0.5% Sodium deoxycholate, 0.5% NP40, 1 mM EDTA, 0.5 mM EGTA), and twice in Wash Buffer

4 (50 mM Tris pH 8.0, 10 mM EDTA, 5 mM EGTA). Chromatin was eluted twice for a total of 30 mins at 65°C with shaking in Elution Buffer (1% SDS, 0.1 M $NaHCO_3$) and eluates were pooled. NaCl was added to samples and inputs to a final concentration of 200 mM, before incubating overnight at 65°C to reverse cross-links. DNA was purifying using the QIAquick PCR Purification Kit (Qiagen).

Table 2.6 List of antibodies used in ChIP experiments

| Target | Species | Antibody | Supplier |
|---|---|---|---|
| H3K27ac | Rabbit | ab4729 | Abcam |
| Nanog | Rabbit | A300-397A | Bethyl Laboratories |
| Sox2 | Rat | 14-9811 | Ebioscience |
| Normal rat IgG | Rat | sc-2026 | Santa Cruz |
| Normal rabbit IgG | Rabbit | sc-2027 | Santa Cruz |

### 2.2.5. Protein extraction and Western blotting

Cells were lysed in RIPA buffer (50 mM Tris-Cl pH8.0, 150 mM NaCl, 1% Tx100, 0.1% SDS, 1mM EDTA) containing Complete-ULTRA protease and PhoStop phosphatase inhibitor cocktails (Roche), and sonicated for 3 cycles, 30 seconds ON 30 seconds OFF in a Bioruptor200 (Diagenode) at high intensity. Debris was removed by pelleting, and aliquots of supernatant were frozen and stored at -80°C or mixed with 4X Bolt LDS Sample Buffer and 10X Bolt Sample Reducing Agent (ThermoFisher). Samples were boiled at 95°C for 5 minutes. SDS-PAGE electrophoresis was performed using Bolt 10% Bis-Tris Plus gels (ThermoFisher) in a Mini Gel Tank (ThermoFisher). Protein transfer was performed using the iBlot2 dry blotting system (ThermoFisher) and iBlot2 nitrocellulose Transfer Stacks (ThermoFisher). Membranes were incubated with primary antibodies (see below). Detection was achieved using HRP-conjugated secondary antibodies against the appropriate species (GE Healthcare) and ECL Plus Western Blotting Detection System (GE Healthcare).

Table 2.7 Antibodies used for Western Blotting

| Target | ID | Supplier |
|---|---|---|
| Oct4 (used in 4.2.1 and 4.2.2) | sc8628 | Santa Cruz Biotechnology |
| Oct4 (used in 4.2.3) | 83932 | Cell Signaling Technology |
| Nanog | A300-397A | Bethyl Laboratories |
| Sox2 | 14-9811 | ThermoFisher |
| α-tubulin | ab7291 | Abcam |

### 2.2.6. Histology

E8.5 embryos were fixed for 4 hours at 4°C with 4% paraformaldehyde (Sigma) in PBS, gradually adjusted to 20% sucrose over 2 days, then mounted in OCT and snap frozen on liquid nitrogen. 8 μm

cryosections were taken and stored at -80°C. Following rehydration in PBS, sections were permeabilised in 0.25% Triton X-100 (Sigma) in PBS, then blocked in 5% donkey serum (Sigma) and 0.1% Triton X-100 in PBS. Sections were incubated overnight at 4°C in blocking buffer with the following primary antibodies: Nanog (1:300, rat mAb, eBioscience); Oct4 (1:300, rabbit mAb, Cell Signaling); Oct4 (1:300, goat pAb, Santa Cruz); Sox1 (1:300, rabbit pAb, Cell Signaling); Sox2 (1:300, rat mAb, eBioscience). The following day, washes were performed with 0.1% Triton X-100 in TBS, and samples incubated with DAPI and AlexaFluor secondary antibodies against the appropriate species at 1:1000 (Life Technologies).

Sections were imaged using a Zeiss ApoTome microscope at 20x then tiled. After imaging, H&E histological staining was performed on the same sections according to standard methodologies. These sections were then re-imaged in the same pipeline. Cyrosectioning and H&E staining were performed by the W-MRC CSCI histology facility, immunochemistry and imaging was performed by Hannah Stuart, as noted in the text.

## 2.3. Bioinformatic analysis

### 2.3.1. ChIP-seq and RNA-seq alignment and visualization

RNA-seq reads were adaptor-trimmed with cutadapt v1.10 (Martin, 2011) via TrimGalore v0.4.1 (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore) and mapped to the mouse reference genome (GRCm38/mm10) with TopHat2 v2.1.0 (https://ccb.jhu.edu/software/tophat) using default parameters for paired end reads. Transcript counts were normalized by library size or by the normalization factor calculated by DESeq2 v1.14.1 (Love et al., 2014) as indicated. ChIP-seq reads were adaptor-trimmed with cutadapt v1.10 via TrimGalore v0.4.1 and mapped to the mouse reference genome (GRCm38/mm10) with BWA v0.7.12-r1039 using the BWA-backtrack algorithm (Li and Durbin, 2009) by calling aln and samse using default parameters for single-end reads. ChIP-seq counts were normalized by library size and extended to match the fragment size. Files were converted to bedgraph format and uploaded to the UCSC Genome Browser for visual inspection. Strand specific RNA-seq files were generated by Michael Barber and Sabine Dietmann using featureCounts (Liao et al., 2014). Gene-wise counts were generated using featureCounts (Liao et al., 2014) to annotate the genes from Ensembl GRCm38.86 release.

### 2.3.2. RNA-seq statistical analysis

Normalization and statistical analyses were performed using DESeq2 v1.14.1 (Love et al., 2014)

For $Oct4^{F/-}$ ESC dataset, two biological replicates of each timepoint were sequenced on two lanes. Technical replicates were collapsed after validating that correlation between technical replicates was greater than correlation across samples. Size factor and dispersal estimation, and model fitting and testing were performed by calling DESeq (DESeq2 v1.14.1) with default parameters. MA plots and PCA analysis were performed using DESeq2.

### 2.3.3. ChIP-seq analysis

Nanog ChIP-seq was performed on duplicate IPs for two hour IAA induced and uninduced samples. Nanog peaks were obtained using MACS2 v2.1.1.20160309 on merged reads of induced samples. Oct4 peaks were obtained using MACS2 on publically available Oct4 ChIP-seq data (GEO accession GSM307140) against whole cell extract control (GEO accession GSM307155) (Marson et al., 2008). Remaining false-positive peaks located in repetitive regions were blacklisted. Differential binding analysis was performed using DiffBind v2.2.12 (Rory and Brown, 2011; Ross-Innes et al., 2012). Normalization factors were calculated using DESeq2. Heatmap and summary plots were generated using DeepTools3 Galaxy Version 3.1.2.0.0 (Ramírez et al., 2016).

# Chapter 3: Establishing the reprogramming capacity of fixed-Oct4 cells after *in vivo* differentiation

## 3.1. Introduction

It is well established that Oct4 plays a vital role in the establishment and maintenance of naïve pluripotency *in vivo* (Le Bin et al., 2014; Nichols et al., 1998), but in recent years it has become increasingly clear that a specific level of Oct4 is critical both for entry into and exit from pluripotency *in vitro*. Radzisheuskaya *et al.* demonstrated that cells successfully completing somatic cell reprogramming to form iPSCs converge to a wild-type level of Oct4 even if it is exogenously expressed (Radzisheuskaya et al., 2013), indicating that a precise level of Oct4 is required at some point in the reprogramming process. Meanwhile, reducing the level of Oct4 in naïve cells while maintaining a basal level of expression (~1/5 to 1/2 a wild-type level) compromises the ability of cells to exit pluripotency (Karwacki-Neisius et al., 2013; Radzisheuskaya et al., 2013). With this in mind, I was interested to see how constitutive expression of an ESC level of Oct4 would affect differentiation *in vivo* and subsequent somatic cell reprogramming.

Specifically, I wanted to examine the capacity of cells maintaining an ESC level of Oct4 to exit pluripotency *in vivo*. Typically, a transient increase in Oct4 expression is observed in differentiation from ESCs (Kalkan et al., 2017), and it is unclear whether this is required for normal differentiation. In the absence of an endogenous allele cells cannot regulate Oct4 in this manner, allowing us to test whether cells can differentiate normally while sustaining an ESC level of expression. Furthermore, given that an ESC level of Oct4 appears to be required at some stage in the reprogramming, I wanted to test whether somatic cells expressing Oct4 at this level throughout the process would demonstrate efficient reprogramming.

### 3.1.1. Generating cells with a fixed ESC level of Oct4

An iPSC line has previously been generated in our group wherein the endogenous Oct4 gene is knocked out from both alleles and Oct4 is expressed entirely from randomly integrated, constitutively expressed Oct4-2a-Cherry transgenes (fixed-Oct4 iPSCs) (Radzisheuskaya et al., 2013, Figure 3.1a). Radzisheuskaya et al. demonstrated that during reprogramming cells select for an ESC level of Oct4; consequently this line expresses a wild-type level of Oct4, and this level is roughly maintained during somatic differentiation (Radzisheuskaya et al., 2013).

### 3.1.2. Cells with an ESC level of Oct4 differentiate and reprogram efficiently *in vitro*

Recent work in our lab has shown that cells maintaining an ESC level of Oct4 can be differentiated to EpiSCs *in vitro* (Stuart et al., in preparation). This demonstrates that they retain the capacity to exit the naïve state; however, it does not resolve whether they are able to make more advanced cell fate decisions, exit pluripotency or whether they differentiate in response to normal cues *in vivo*. Nonetheless, it has been exciting to observe that these fixed-Oct4 EpiSCs are able to revert to naïve pluripotency solely by switching to naïve growth conditions (Figure 3.1b) with remarkable speed and efficiency, and without requiring induction of naïve-specific transcription factors or application of chromatin-modifying factors. It is therefore interesting to investigate the capacity of fixed-Oct4 iPSCs for more advanced differentiation, and to test the limits of their reprogramming potential.

Figure 3.1

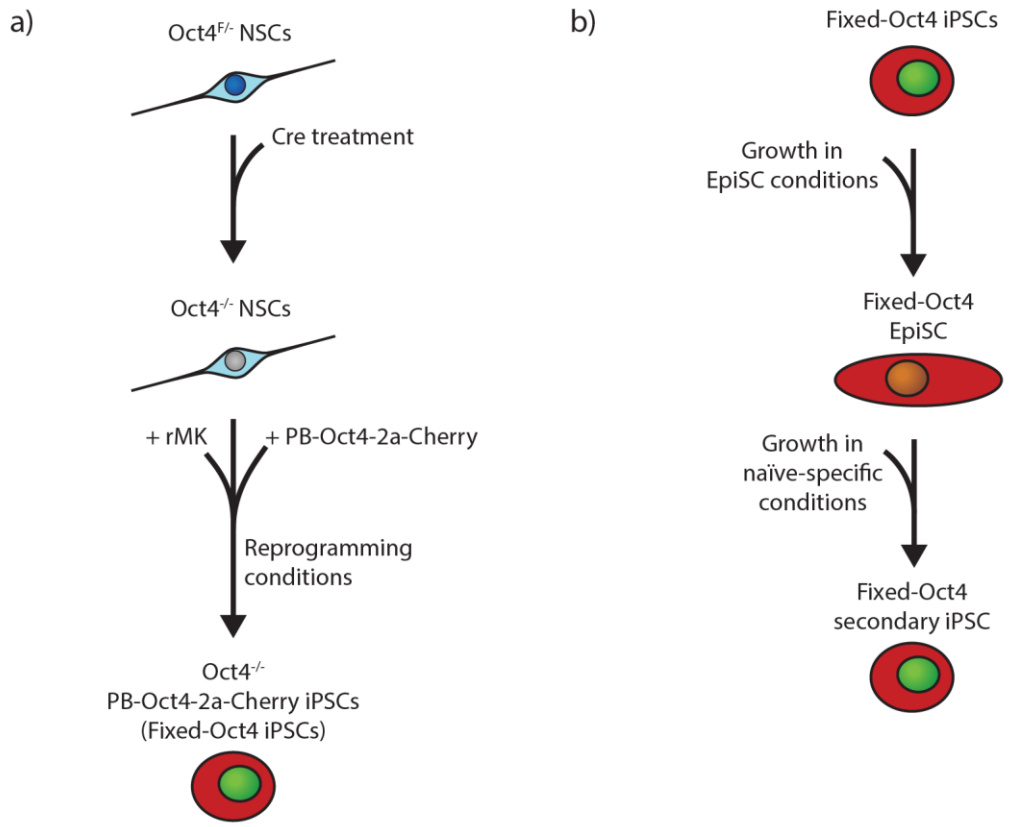



Figure 3.1 Fixed-Oct4 PSCs can be differentiated *in vitro* and revert to naïve pluripotency on media switch
a) Schematic of the protocol used by Aliaksandra Radzisheuskaya to generate iPSCs with constitutive expression of Oct4 at an ESC level (fixed-Oct4 iPSCs). Briefly, NSCs were obtained from mice carrying one knocked-out and one floxed Oct4 allele ($Oct4^{F/-}$ NSCs). These were treated with Cre to generate Oct4-null NSCs ($Oct4^{-/-}$ NSCs). Cells were transduced with retroviral cMyc and Klf4 (rMK) and transfected with a randomly integrating, constitutively expressed Oct4-2a-Cherry construct (PB-Oct4-2a-Cherry). This yielded red fluorescent iPSCs (fixed-Oct4 iPSCs) expressing Oct4 at an ESC level, which is largely maintained during differentiation. b) Schematic of experiments carried out by Hannah Stuart demonstrating that fixed-Oct4 iPSCs are capable of *in vitro* differentiation to EpiSCs and subsequent reversion to iPSCs on media switch. Fixed-Oct4 iPSCs generated by Radzisheuskaya were grown in EpiSC maintenance conditions. After a period of adjustment, cells emerged with an EpiSC-like morphology. Transcriptional profiling confirmed that these cells had attained an EpiSC identity, and their complete loss of naïve pluripotency was demonstrated by the inability to contribute to chimeras following blastocyst injection. Changing the exogenous signals to naïve-specific growth conditions was sufficient to revert these cells to an ES-like identity. This was confirmed by transcriptional changes and by acquisition of competency to contribute to chimeras on blastocyst injection. Meanwhile, control EpiSCs could not revert to naïve pluripotency solely on changing the signalling environment, requiring exogenous expression of one of several naïve transcription factors to successfully reprogram.

## 3.2. Results

### 3.2.1. Fixed-Oct4 cells are able to differentiate *in vivo* and integrate well into E8.5 chimeras

In order to produce appropriately differentiated cells, we generated chimeric embryos through blastocyst injection of fixed-Oct4 iPSCs into wild-type blastocysts (Figure 3.2a). By late E8.5, most embryos had completed turning. At this stage, no clear abnormalities were observed from a visual

analysis (Figure 3.2b). Chimeric embryos could not be distinguished from non-chimeric littermates other than by fluorescent analysis of Cherry signal derived from the constitutive Oct4-2a-Cherry transgene. Most embryos showed a high level of chimerism with contribution from fixed-Oct4 cells throughout the embryo (Figure 3.2b and Table 3.1).

Figure 3.2

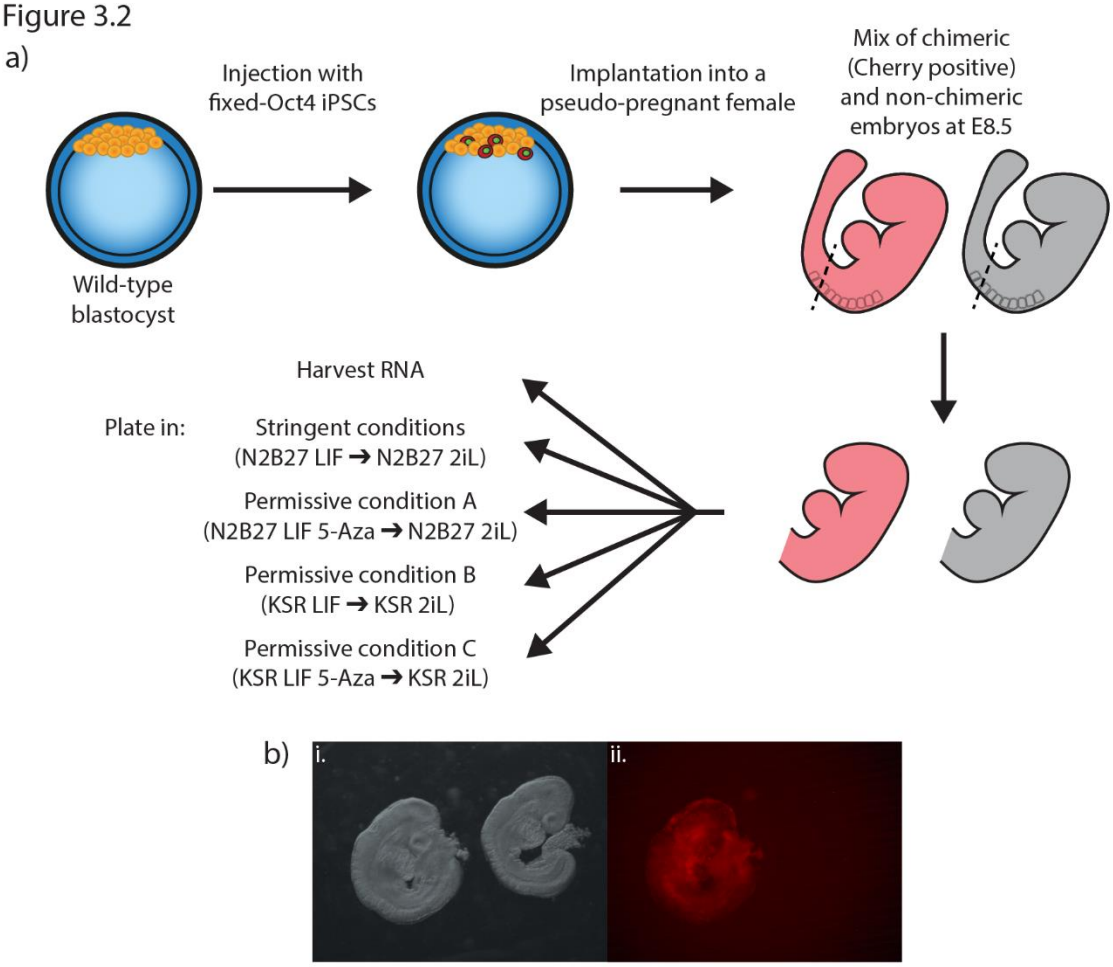



Figure 3.2 Generating fixed-Oct4 chimeric embryos
a) Schematic for generating fixed-Oct4 chimeric embryos and subsequent profiling of their reprogramming capacity. The Zona Pellucida of wild-type early blastocysts was pierced with a laser. Fixed-Oct4 iPSCs were then injected into the blastocoel, and the embryos were transferred to the uterus of a pseudo-pregnant mouse. The blastocysts implant and continue to develop, and injected cells may contribute to the embryo yielding a chimera (detectable due to constitutive Oct4-2a-Cherry expression). The posterior and extraembryonic tissues were removed, and the anterior material was partially dissociated to collect samples for RT-qPCR analysis and to assay for reprogramming capacity. b) i. Phase contrast and ii. fluorescence images of chimeric (left) and non-chimeric (right) late E8.5 embryos prior to removal of the posterior end. Note the extensive contribution of Cherry-positive iPSCs to the chimeric embryo. Blastocyst injection was performed by William Mansfield.

In order to examine the degree of chimeric contribution more thoroughly, and to check that fixed-Oct4 cells had integrated into differentiated tissues, we performed cryosectioning and immunofluorescence staining. Figure 3.3, adapted from the graduate thesis of Hannah Stuart, shows

extensive contribution of Cherry positive cells throughout the E8.5 embryo. Oct4 staining corresponds well with Cherry signal, indicating that Cherry is a good proxy for the presence of Oct4. It should be noted that Nanog is specifically localised to the nucleus in PGCs and pluripotent ESCs and EpiSCs (Chambers et al., 2007; Gillich et al., 2012; Yamaguchi et al., 2005); specific (nuclear) Nanog signal is entirely absent indicating that all cells have successfully exited pluripotency (Figure 3.3a), but also that this section does not contain any germ cells from either parent. Sections of another embryo demonstrated even higher levels of chimerism (Figure 3.3b). In this embryo, it is clear that Oct4 is well integrated into various tissues, with extensive contribution to the myocardium and somites. Anteriorly, there is significant overlap in staining for Sox1 and Sox2 with Oct4, showing that fixed-Oct4 cells are differentiating efficiently into neural lineages, while distal overlap is indicative of gut endoderm differentiation.

Figure 3.3

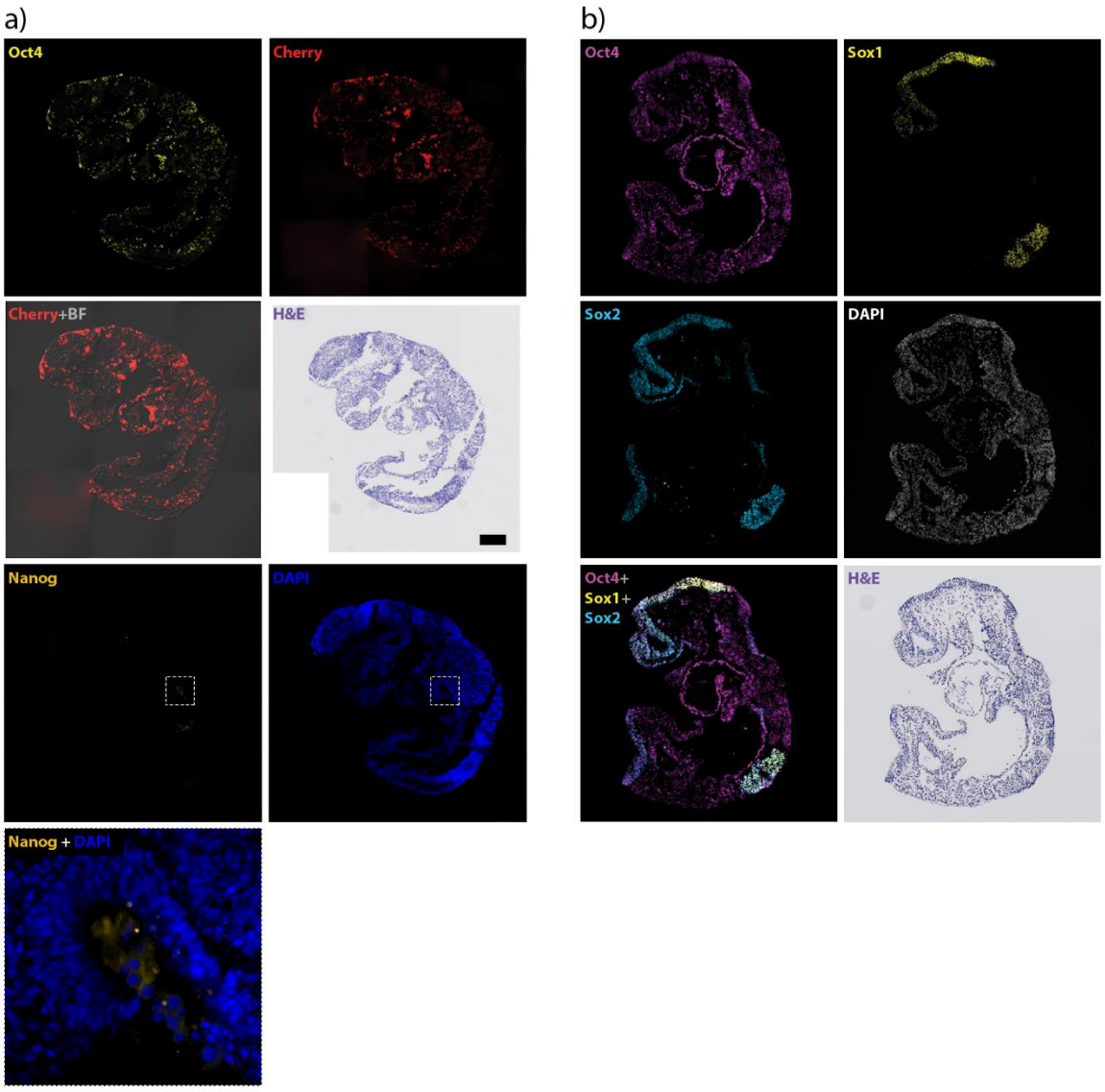



Figure 3.3 Cryosections and immunofluorescence of E8.5 chimeras (data from Hannah Stuart)
8µm cryosections were taken of E8.5 fixed-Oct4 chimeras. Subsequently expression of pluripotency and differentiation markers was assessed by IF, using antibodies against Oct4, Sox1, Sox2 and Nanog as stated. Cherry was detected directly by fluorescence and nuclei were stained with DAPI. H&E staining was performed on the same sections after imaging. a) IF against Oct4 and Nanog, plus imaging of Cherry, DAPI and H&E staining. Oct4 IF shows that Cherry accurately indicates cells in which Oct4 is expressed. Fixed-Oct4 cells have successfully integrated throughout the embryo. Dotted lines indicate inset below; the absence of nuclear Nanog signal demonstrates that all cells have exited pluripotency. Non-nuclear Nanog stain is likely to be unspecific signal. b) IF against Oct4, Sox1 and Sox2, plus imaging of DAPI and H&E staining. Anterior overlapping Oct4, Sox1 and Sox2 signal reveal that fixed-Oct4 cells are capable of ectodermal differentiation *in vivo*, and contribute well to neural lineages, while posterior Oct4 and Sox2 overlap indicate gut endoderm differentiation.
Cryosectioning and H&E staining performed by the SCI histology facility, IF and imaging performed by Hannah Stuart, figure adapted from the dissertation of Hannah Stuart.

### 3.2.2. *In vivo*-differentiated fixed-Oct4 cells reprogram efficiently

Given that these embryos did not show any obvious defects or tumours, we concluded that early differentiation events had occurred relatively normally in these embryos. In order to examine the

reprogramming potential of the differentiated fixed-Oct4 cells, we removed the posterior half of the embryos in order to prevent contamination with any remnants of the primitive streak and associated pluripotent cells, though none should remain (Osorno et al., 2012), and any primordial germ cells or their precursors (Molyneaux et al., 2001; Tam and Snow, 1981). We manually dissociated the anterior section of the embryos to small clumps of cells and plated these in a range of media conditions (Figure 3.2a), reserving some material for RT-qPCR analysis.

Dissociated cells were plated in stringent and permissive reprogramming conditions. There is evidence that immediate application of naïve-specific 2iL conditions to differentiated cells inhibits reprogramming (Silva et al., 2008), an effect that is likely due to inhibition of key survival signalling before reprogramming occurs. Indeed, in preliminary experiments, dissociated anterior cells plated directly in 2iL failed to reprogram, while germ-cell-containing allantois readily generated EGCs. To avoid this inhibition, cells were grown for one week in reprogramming media lacking 2i before switching to 2iL conditions. In permissive conditions cells were maintained for one week in basal media plus LIF in the presence of 5-azacitidine, a DNA methyltransferase inhibitor (Stresemann and Lyko, 2008), and/or ascorbic acid (in the form of KSR-based media (Blaschke et al., 2013)), a Tet enzyme and JmjC-domain histone demethylase agonist (Blaschke et al., 2013; Klose et al., 2006; Yin et al., 2013), both of which lead to a global reduction in DNA methylation (Blaschke et al., 2013; Mikkelsen et al., 2008; Yin et al., 2013) and have been shown to improve both the rate and efficiency of somatic cell reprogramming (Esteban et al., 2010; Mikkelsen et al., 2008; Tsuji-Takayama et al., 2004; Wang et al., 2011). After this, media was changed to N2B27 2iL or KSR 2iL, and after a further 5 days selection for endogenous Oct4 promoter activity was applied. Meanwhile, in stringent conditions cells were maintained in N2B27 LIF prior to media switch to N2B27-2iL and subsequent selection. Medium change to ESC maintenance conditions is reportedly sufficient to revert a fraction of some EpiSC lines to the naïve state (Bao et al., 2009); however, generally reprogramming to naïve pluripotency requires the addition of transgenes, altered signalling environments or epigenetic modifiers (Gillich et al., 2012; Guo et al., 2009; Guo and Smith, 2010; Hanna et al., 2009; Murayama et al., 2015; Silva et al., 2009; Zhang et al., 2016). In this study, cells should be more advanced in their differentiation than EpiSCs, having already undergone the events of gastrulation and lineage segregation, and so we would not anticipate that any wild-type cells would reprogram under these stringent conditions.

Of the 11 embryos collected, 10 were chimeric by visual inspection of Cherry contribution, with one non-chimeric embryo which we used as a negative control (Table 3.1). Additionally, we plated material from the posterior end of the embryos together with allantois tissue; these regions should contain migrating primordial germ cells (Molyneaux et al., 2001; Tam and Snow, 1981), which should readily convert to embryonic germ cells (EGCs) in our reprogramming conditions (Leitch et al., 2010), and thus act as a positive control.

| | Embryo number | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| Chimeric? | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | No | Yes |
| Reprogrammed in stringent conditions? | Yes | No | Yes | Yes | Yes | Yes | Yes | No | No | No | Yes |
| Reprogrammed in permissive conditions? | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | No | Yes |

Table 3.1 List of E8.5 embryos harvested showing their chimerism and reprogramming capacity
Chimeras are listed as chimeric if there was observable red fluorescent representing the Oct4-2a-Cherry transgene in the embryo prior to dissociation. This was confirmed by monitoring Cherry fluorescence in plated cells and by continued expression of Oct4 as detected by RT-qPCR on dissociated embryonic tissue. Successful reprogramming was determined by the establishment of cell lines that could be serially passaged and showed dome-shaped morphology, activation of the endogenous Oct4 locus, and acquired expression of pluripotency markers Nanog and Rex1 by RT-qPCR analysis.

Figure 3.4

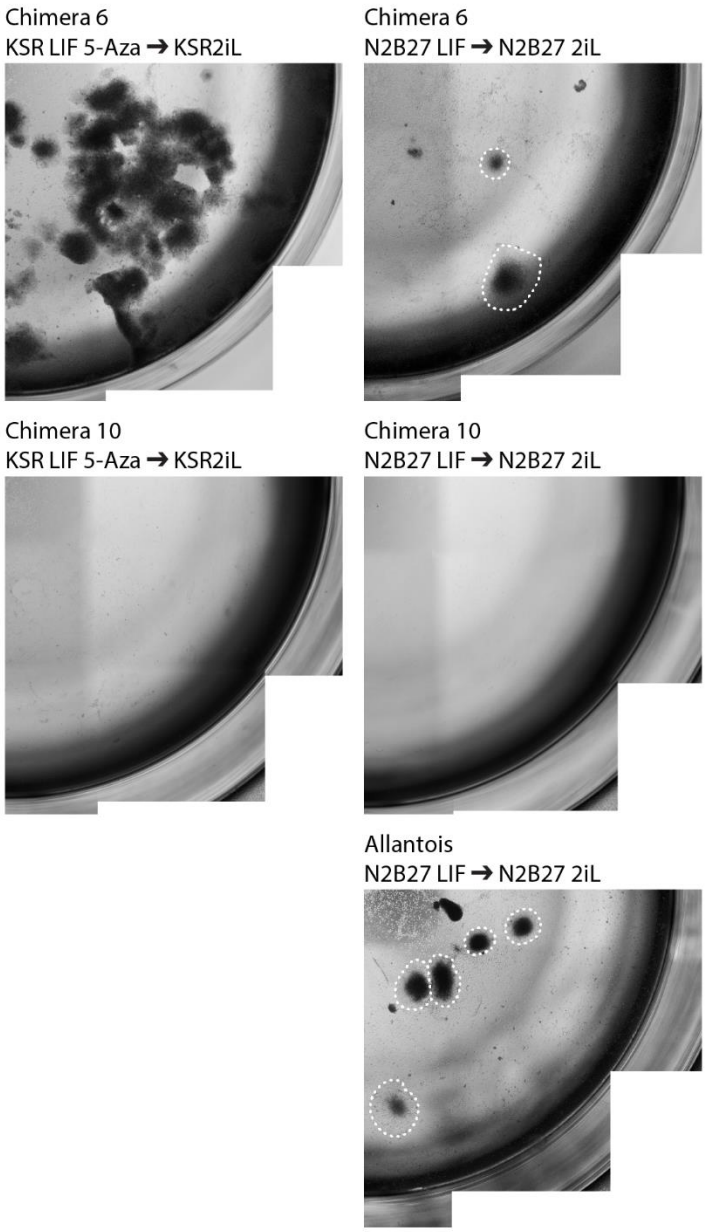



Figure 3.4 Emerging iPS colonies from fixed-Oct4 cells differentiated *in vivo*
Phase contrast images of a quarter of a 24 well of emerging iPSC colonies from embryo 6 (top) which is representative of chimeric embryos, embryo 10 (middle) which was not chimeric by visual inspection of Cherry fluorescence, and chimeric allantois (bottom) in permissive conditions (left) or stringent conditions (right). Stringent conditions: N2B27 LIF → N2B27 2iL; permissive conditions: KSR LIF 5-Aza → KSR 2iL. Colony boundaries are marked with a dotted line in stringent conditions; in permissive conditions, efficiency was high and cell grew quickly so the majority of colonies have merged making it difficult to define their edges. Images were taken prior to passaging; note these colonies are overgrown as wells were maintained and passaged in parallel. The majority of non-reprogrammed material has been cleared, and colonies are expanding well.

As detailed in Table 3.1, all of the chimeric embryos were able to generate iPS lines in the permissive conditions, typically with very high efficiencies (Figure 3.4, top left), while the non-chimeric control did not (Figure 3.4, middle left). This implies that constitutive expression of Oct4 at an ESC level instils competency to post-implantation somatic cells to reprogram to the naïve state. More remarkably, we

established iPS lines from 7 of the 10 chimeric embryos (70%) under our stringent reprogramming conditions (Figure 3.4, top right), relying solely upon exogenous Oct4 expression and a change in the signalling environment to revert their identity to naïve pluripotency. This is particularly noteworthy given the small amount of starting material; cell numbers were limited since the posterior end of the embryo was removed, a portion was retained for expression analysis and the remainder was split into four reprogramming conditions. Consequently each reprogramming well contained approximately 10% of a single embryo, equating to perhaps $5x10^5$ cells (Cho et al., 2011).

Figure 3.5

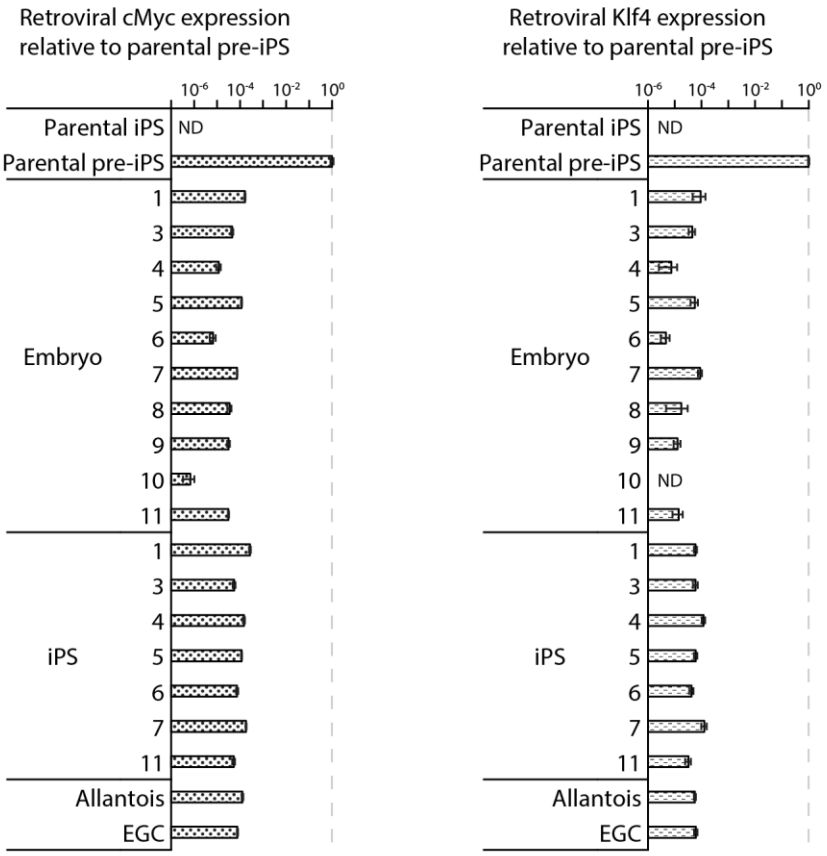



Figure 3.5 Retroviral transgenes remain silent during fixed-Oct4 reprogramming
RT-qPCR analysis of parental iPS and pre-iPSCs, material extracted from E8.5 embryos at the time of plating, established iPS lines from material plated in stringent conditions and passaged once since derivation, allantois material from E8.5 embryos and an established EGC line from plated allantois, as indicated. ND: Not determined; expression was below the threshold of detection for some samples: parental iPS rKlf4 < $1.3x10^{-4}$, rcMyc < $6.7x10^{-5}$; embryo 10 rKlf4 < 2.0x10-7. Values represent the mean of three replicate qPCR reactions and error bars represent ± SD; values were calculated by the ΔΔCt method using the housekeeping gene Gapdh as a reference gene and normalized to parental pre-iPSCs. Expression is shown on a $log_{10}$ scale.

The fixed-Oct4 iPSCs were generated using retroviruses which are stably silenced during the reprogramming process. In order to validate that Oct4 was the sole reprogramming factor present, I performed RT-qPCR using probes specific for the retroviral Klf4 and cMyc transgenes. The transgenes were not reactivated in the embryonic tissue (Figure 3.5).

Figure 3.6

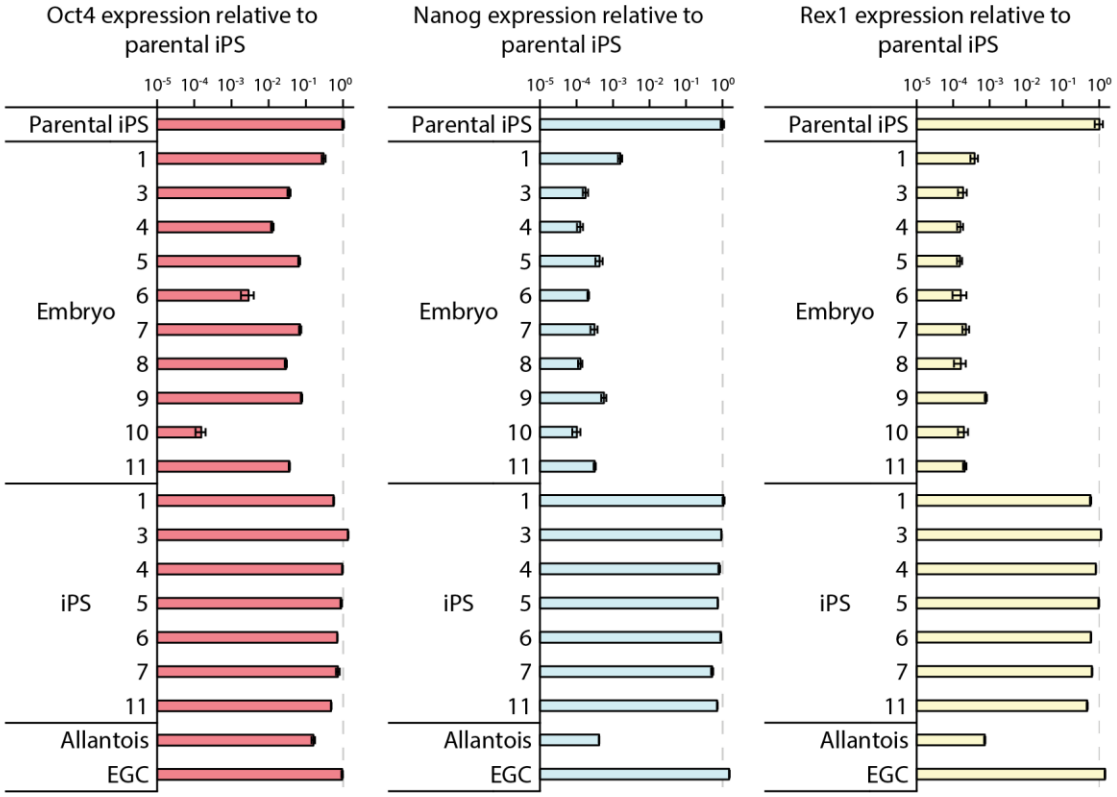



Figure 3.6 Analysis of expression of pluripotency genes in chimeric embryos and iPSCs
RT-qPCR analysis of parental iPSCs, material extracted from E8.5 embryos at the time of plating, established iPS lines from material plated in stringent conditions and passaged once since derivation, allantois material from E8.5 embryos and an established EGC line from plated allantois, as indicated. Values represent the mean of three replicate qPCR reactions and error bars represent ± SD; values were calculated by the ΔΔCt method using the housekeeping gene Gapdh as a reference gene and normalized to parental iPSCs. Expression is shown on a $\log_{10}$ scale. Oct4 was detectable at variable levels in all samples, though embryo 10 (Cherry negative) had an order of magnitude less than the next lowest embryo. Nanog and Rex1 levels were lower in the embryonic tissue by several orders of magnitude compared to the reference iPSCs, implying complete silencing. iPSCs and EGCs showed Nanog and Rex1 expression at levels comparable to parental iPSCs.

It should be noted that the cells that we plated do appear to have fully exited pluripotency *in vivo*, as expression of the naïve-specific marker Rex1 and the general pluripotency marker Nanog were extinguished in the chimeric embryos as well as in the non-chimeric control (Figure 3.6). Both markers are expressed robustly in the successfully reprogrammed iPSCs at levels comparable to the parental iPSCs and the positive control EGCs (Figure 3.6).

### 3.2.3. Constitutive Oct4 expression causes several severe malformations in E12.5 chimeras

Given that cells with an ESC level of Oct4 appear to undergo normal early embryogenesis, we decided to examine their capacity to contribute to more advanced embryos (Figure 3.7a). Extensive chimerism was observed in E12.5 embryos, with contribution to organs and somites, and throughout the body

(Figure 3.7b and c). However, this was accompanied by several developmental defects with varying levels of severity between chimeras (Table 3.2).

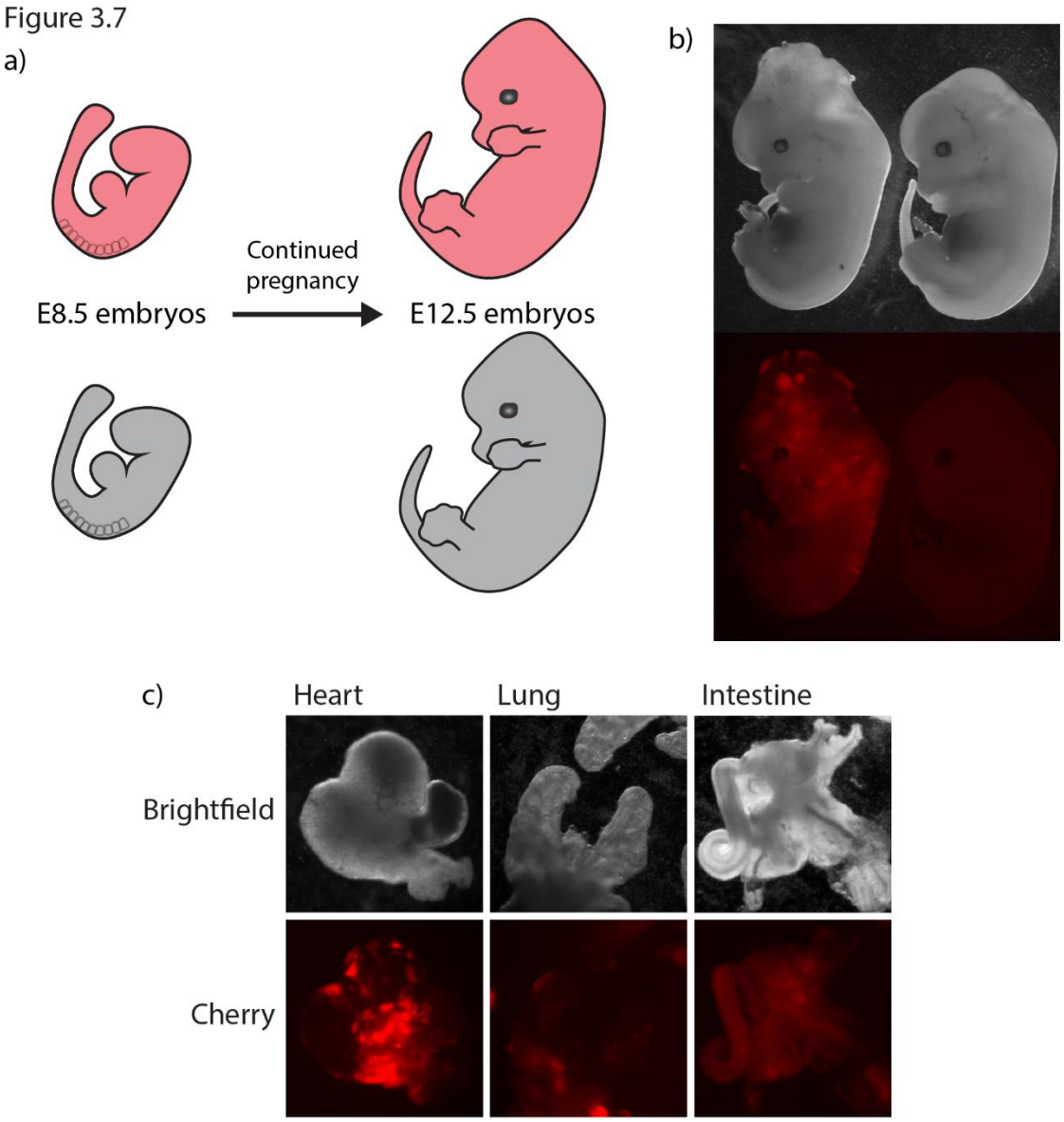



Figure 3.7 Fixed-Oct4 cells contribute efficiently to E12.5 embryos
a) Schematic showing the possible outcome of continued *in vivo* development yielding relatively unperturbed E12.5 chimeric and non-chimeric embryos. b) Images of E12.5 chimeric (left) and non-chimeric (right) embryos, phase contrast (top) and Cherry (bottom). Significant outward expansion of the midbrain in chimeric embryos is clearly visible, along with large fluorescence-positive growths. Fixed-Oct4 cell contribution can be observed throughout the embryo. c) Phase contrast and Cherry images of heart, lung and intestine from chimeric embryos showing patches of fixed-Oct4 cells in developing organs. It should be noted that non-chimeric intestine displayed significant autofluorescence, so level of contribution is unclear.

An immediately obvious phenotype was an extreme expansion of the mesencephalon. In over one third of chimeras (Table 3.2) the midbrain shows sufficient growth in the confined space that it buckles and pushes outwards (Figure 3.7b). It is impossible to ignore that the midbrain shows a large number

of Oct4 Sox1 double positive cells in our E8.5 staining (Figure 3.3b). Sox1 is a neurepithelial determinant which maintains neuronal precursors in a primitive state, with low expression of glial markers BLBP and β3-tubulin (Suter et al., 2009). Cells lose Sox1 expression as they differentiate (Pevny et al., 1998), and overexpression of Sox1 during differentiation leads to expansion of the self-renewing neurepithelial population (Suter et al., 2009). It is therefore tempting to speculate that an ESC level of Oct4 disrupts the normal loss of Sox1 expression in the midbrain, causing an increase in proliferation leading to this phenotype. However, this would require more investigation to confirm. Despite this, we were able to derive NSCs from such chimeric embryos with no obvious phenotype at early passage numbers, though these have not been extensively characterized. Additionally, a small number of embryos displayed exencephaly (Figure 3.8a and Table 3.2), in which the brain did not appear to be fully covered by surface ectoderm. Exencephaly normally results from an earlier failure to fully close the neural tube (Macdonald et al., 1989; Wallace et al., 1978). Interestingly, exencephaly commonly displays partial penetrance, and often co-occurs with blebs in the epidermis (Wallace et al., 1978) which we observed in over half of the chimeras (Table 3.2). Alternatively, it has been previously noted that overexpansion of the midbrain can lead to exencephaly-like abnormalities; embryos in which RhoA is knocked out in the mesencephalon demonstrate midbrain expansion similar to our embryos and subsequently the brain protrudes, without any sign of failure to close the neural tube (Katayama et al., 2011).

| 1st round | Embryo | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | Summary |
| Skin blebs | Y | Y | N | Y | N | N | Y | Y | 5 |
| Vascular buds | Y | N | Y | N | Y | Y | N | Y | 5 |
| Malformed forelimbs | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 individuals |
| Malformed hindlimbs | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 individual |
| Midbrain expansion | N | N | Y | N | Y | Y | N | N | 3 |
| Exencephaly | N | N | N | N | Y | N | N | N | 1 |

| 2nd round | Embryo | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7† | 8 | 9 | 10 | Summary |
| Skin blebs | Y | N | N | Y | Y | Y | Y | Y | Y | Y | 8 |
| Vascular buds | Y | Y | N | Y | N | Y | N | N | Y | N | 5 |
| Malformed forelimbs | 1 | 2 | 0 | 0 | 2 | 2 | 2 | 1 | 0 | 0 | 6 individuals |
| Malformed hindlimbs | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 individual |
| Midbrain expansion | N | Y | Y | Y | Y | N | N | N | N | N | 4 |
| Exencephaly | N | Y | N | N | Y* | N | N | N | N | N | 2 |

| | Number of chimeras | Deformity | | | | | |
|---|---|---|---|---|---|---|---|
| | | Skin blebs | Blood sacs | Malformed limbs | Midbrain expansion | Exencephaly | Other |
| 1st round | 8 | 5 | 5 | 1 | 3 | 1 | 0 |
| 2nd round | 10 | 8 | 5 | 6 | 4 | 2 | 1 |
| Total | 18 | 11 (61%) | 10 (56%) | 7 (39%) | 7 (39%) | 3 (17%) | 1 (6%) |

Table 3.2 Score of malformations by embryo and summary of malformations
E12.5 embryos were harvested on two occasions following blastocyst injection with 1-5 ESCs and transfer to pseudo-pregnant females. Embryos were imaged prior to dissection for derivation of MEF and NS lines or fixation for later analysis, and these images were examined for malformations. Note, limb malformations were not noticed during harvesting or imaging of the first round of embryos; consequently images were typically taken of one side, obscuring two limbs. As a result, the number of embryos harbouring limb malformations may be higher than reported for the first round. Brain sizes were not measured due to the fact that embryos appeared to be at slightly different stages resulting in variation in their overall size. Instead, midbrain expansion was counted if the overall structure of the midbrain was obviously protruding out from the normal brain cavity while remaining covered by surface ectoderm.
* Embryo 5 displayed partial exencephaly; only the hindbrain and the rear of the midbrain were exposed. † Embryo 7 displayed an obviously underdeveloped lower jaw.

Figure 3.8

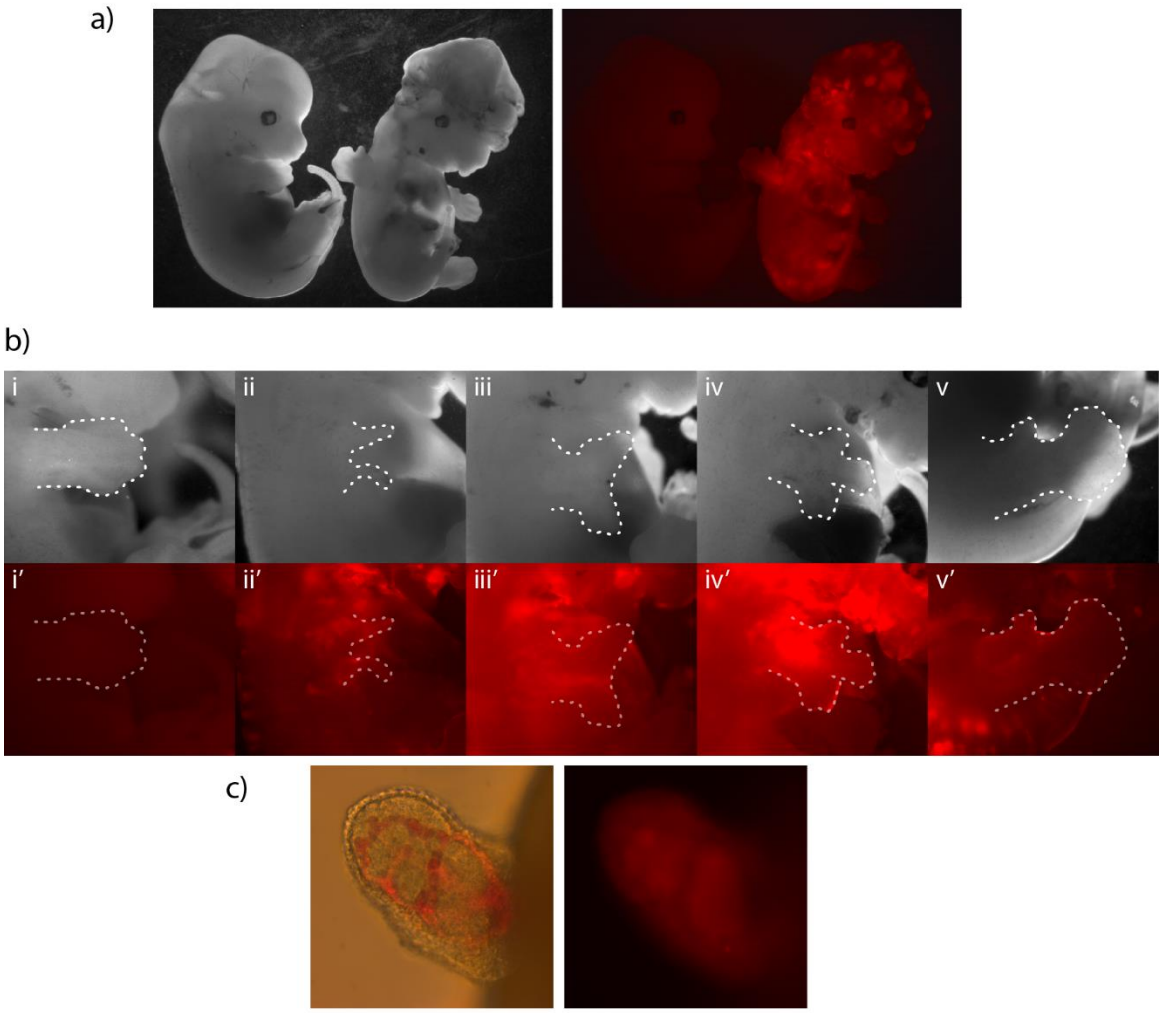



Figure 3.8 E12.5 fixed-Oct4 chimeric embryos show a range of defects, often with partial penetrance
a) Phase contrast and fluorescence images of a normal, non-chimeric E12.5 embryo (left), and the most exencephalic chimeric embryo (right). Note that this chimeric embryo also displays malformed forelimbs. b) Images of limbs of E12.5 embryos. i-v: phase contrast images; i′-v′: fluorescence images. i-iv and I′-iv′: forelimbs, v and v′: hindlimb. i and i′ are of a non-chimeric embryo and represent normal limb structure at E12.5. c) Brightfield (left) and fluorescence (right) images of a vascularized bud structure. Note that the red colouring in the brightfield image is intrinsic to the cells and not caused by fluorescence.

Roughly one third of chimeras demonstrated dramatic limb malformations with various levels of penetrance (Figure 3.8b), though only one or two limbs were ever affected (Table 3.2). Some of these malformations (Figure 3.8b ii. and iii.) resemble the phenotype observed in classical transplantation experiments involving engraftment of an additional zone of polarizing activity (ZPA) to the anterior side of the limb bud. This results in an anterior-posterior (A-P) mirroring of limb formation (Saunders, J. W., 1968; Tickle et al., 1975) as the limb A-P axis is established by a gradient of SHH secreted by the ZPA (Riddle et al., 1993). This is a striking phenotype, normally only seen on ZPA transplantation or exogenous application of SHH or retinoic acid (López-Martínez et al., 1995;

Summerbell, 1983; Tickle et al., 1982). It would be interesting to investigate the mechanism behind this phenotype further; in particular future work could elucidate whether constitutive Oct4 expression results in aberrant formation of a second ZPA at the anterior side of the limb bud or if it affects the way cells interpret the normal SHH gradient. Unfortunately, this is beyond the scope of this thesis, but such work could uncover interesting details relating to the combinatorial manner in which local signalling environments interact with transcriptional networks to make cell fate and patterning decisions in development.

The final phenotype we observed was the generation of small, vascularized buds of tissue on the exterior surface of roughly half of the chimeric embryos (Table 3.2). These structure typically contained blood within a complex structure of branched vessels (Figure 3.8c). It is not clear what these structures represent, nor whether they are connected to the embryonic vasculature or if they are a site of haematopoiesis. They were most often found growing in around the mouth.

## 3.3. Discussion

By working with cells constitutively expressing Oct4 at roughly an ESC level, we have demonstrated that Oct4, in combination with naïve-specifying environmental conditions, is sufficient to efficiently reprogram somatic cells to an iPSC identity. The property of reprogramming iPSCs to self-select for an ESC level of Oct4 (Radzisheuskaya et al., 2013) makes it trivial to generate a naïve cell line constitutively expressing Oct4 at this level from a transgene. Using this reagent we have confirmed the results of Radzisheuskaya et al. (Radzisheuskaya et al., 2013) that continued expression of Oct4 is compatible with exit from pluripotency (Figure 3.3 and Figure 3.6). Oct4 overexpression is sufficient to induce differentiation of naïve pluripotent stem cells *in vitro* (Niwa et al., 2000), so it is not entirely surprising that maintenance of Oct4 expression is compatible with differentiation *in vivo*; however, given the level-specific phenotypes that Oct4 demonstrates, with reduced Oct4 expression compromising differentiation both *in vitro* (Karwacki-Neisius et al., 2013; Radzisheuskaya et al., 2013) and *in vivo* (Radzisheuskaya et al., 2013), it is valuable to validate this finding. Notably, differentiation of pluripotent cells constitutively expressing Oct4 has been demonstrated in other systems *in vitro* and *in vivo*, but in these cases the transgene level was known to be significantly lower (Niwa et al., 2000; Ramos-Mejía et al., 2005) or far greater (Niwa et al., 2000) than that seen in ESCs. Additionally, it was introduced into cells already expressing endogenous Oct4, likely resulting in selection for greatly reduced transgene levels.

However, Radzisheuskaya *et al.* did not investigate the reprogramming capacity of differentiated fixed-Oct4 cells. It has previously been shown that overexpression of Oct4 can lead to reprogramming of somatic cells, especially under permissive conditions, though this typically takes 3-6 weeks (Kim et al., 2009; Li et al., 2011; Tsai et al., 2011; Yuan et al., 2011). In our permissive reprogramming experiments, cells were maintained in basal media containing ascorbic acid, 5-azacitidine and LIF for one week before switching to naïve growth conditions (supplementing the basal media with 2i); after five days we selected for activation of the endogenous Oct4 locus, and colonies emerged from all chimeric embryos. This compares favourably with Oct4-induced reprogramming in the presence of multiple chemical reprogramming factors (Li et al., 2011; Yuan et al., 2011).

However, it is remarkable that we observed the generation of iPSCs in the same short timespan in our restrictive reprogramming conditions; we cultured cells for one week in basal media with LIF, lacking any naïve specifiers or epigenetic modifiers, and again when we switched to naïve culture media we obtained lines with activation of the endogenous Oct4 locus within five days from 70% of our embryos (Table 3.1). In earlier experiments (not shown), we had assumed that addition of small molecules such as ascorbic acid and 5-azacitidine would be necessary to overcome the epigenetic barriers established during differentiation, and that Oct4 alone would only be able to reprogram a small fraction of cells over a protracted period in the absence of these factors, as has previously been observed (Kim et al., 2009; Tsai et al., 2011). Instead, it appears that when Oct4 is present at an ESC level it is capable of rapidly reverting cells to naïve pluripotency, requiring less than two weeks as opposed to the 4-5 weeks reported following retroviral transduction of Oct4 (Kim et al., 2009).

It is important to note, however, that the constitutive expression of Oct4 could have affected the transcriptional and epigenetic state of these cells. This might have generated a state in which cells were more prone to reprogram. This is difficult to test, as it is not straightforward to allow cells to differentiate normally with endogenous gene expression, then remove the endogenous allele and transgenically express the gene at that specific level. An additional factor is that we are utilizing a secondary system, reprogramming cells that have already been reprogrammed once before. Typically, secondary systems have far greater reprogramming efficiencies than primary systems (Nagy, 2013). Ideally, to gauge the relative efficiency of reprogramming of fixed-Oct4 cells they should be directly compared to iPSCs that are wild-type with respect to Oct4 to control for this. This sort of experiment would also require accurate counts of starting cells, ideally isolated through FACS sorting to remove

wild-type host cells, and counts of generated colonies. In this way we could conclusively determine whether the fixed-Oct4 cells have an advantage in reprogramming.

It should be noted that cells required a period in media lacking 2i. While PGCs from the allantois reprogrammed efficiently to EGCs on direct application of N2B27-2iL, similar treatment did not result in the generation of any iPSC colonies from the anterior segment of our chimeras. This is indicative that they did not contain a small population of cells retaining pluripotency or any teratocarcinomas, which would be capable of outgrowth in 2iL conditions.

Interestingly, it has previously been reported that cells of the post-implantation epiblast lose the capacity to form primed pluripotent EpiSC cultures between E7.5 and E8.25, but exogenous Oct4 expression extends this window to E14.5, albeit with significantly reduced efficiency at later times (Osorno et al., 2012). In the work presented in this chapter, we dramatically expand on this, demonstrating that maintaining an ESC level of Oct4 with simple naïve-specifying culture conditions is sufficient to reprogram *in vivo* differentiated cells to naïve pluripotency.

We observed a range of interesting phenotypes in E12.5 chimeras (Table 3.2). In normal development, Oct4 is downregulated in most somatic cells during and soon after gastrulation (Osorno et al., 2012). However, some expression is observed broadly in the gastrulating embryo, particularly in the mesoderm until the 8 somite stage, at which point it becomes restricted to the germ lineage (Downs, 2008). Constitutive ubiquitous Oct4 expression in embryogenesis has been studied before (Ramos-Mejía et al., 2005), using a CAG promoter to drive Oct4 expression in all cells of the embryo. There are similarities and differences in the developmental abnormalities we observe, which may relate to differences in Oct4 expression level, or non-cell-autonomous interactions with the wild-type cells present in our chimeras.

At E9.5, Ramos-Mejía *et al.* identified a reduction in the size of the forebrain and midbrain due to increased cell death in Oct4-expressing embryos, with recovery to normal brain morphology by E12.5. The effect of Oct4 on the brain is dramatically different in our chimeric embryos. Although we have not quantitated brain sizes, in part due to the difficulty in identifying identically staged embryos, there was no obvious difference in the scale of the forebrain or midbrain between wild-type embryos and fixed-Oct4 chimeras at E8.5. However, at E12.5 many of our chimeric embryos showed gross expansion of the midbrain (Figure 3.8b and Table 3.2). It is interesting to note that deletion of Oct4 in

the post-implantation embryo at around E7.5 results in a diminished midbrain and greatly reduces cranial mesenchyme density (DeVeale et al., 2013); however, it was suggested that this was due to reduced proliferation in the primitive streak, where Oct4 expression normally persists at this time. This suggests that either expansion of the primitive streak continues for a longer time in our chimeric embryos, or that the expansion of the midbrain that we observe is due to a different mechanism. Furthermore, we observed a small number of exencephalic embryos at E12.5 (Table 3.2). While the exact cause here is unclear, exencephaly is typically the result of failure in anterior neural tube closure (Harris and Juriloff, 2007; Macdonald et al., 1989; Wallace et al., 1978); this could be related to the defective closure of the posterior neural tube observed by Ramos-Mejía *et al.* (Ramos-Mejía et al., 2005).

Ramos-Mejía *et al.* also found that E12.5 embryos were smaller overall, with failed development of digits and zeugopod elements in the limbs, a kinked tail, and a clear reduction in the mandibular arch resulting in a small lower jaw. The fixed-Oct4 chimeric embryos did not appear any smaller than their non-chimeric littermates. We did observe a dramatic phenotype in the limbs, similar to that seen previously. However, we noted highly variable penetrance, with many embryos displaying three normal limbs and a single defective limb, and a particular prevalence in forelimbs (Table 3.2). The defect also appeared to be variable, with some embryos appearing to duplicate the limb axis and others generating secondary limb outgrowths. We did not examine the tail in detail, but there was no obvious sign of posterior failure to close the neural tube, and no gross increase in the number of somites (Aires et al., 2016). Similarly, the structure of the jaw was not examined thoroughly, but was clearly greatly reduced in a single chimera (Table 3.2).

It is also interesting to compare our findings to the effects of ectopic Oct4 expression in adult mice. Previously, ES-derived and chimeric mice carrying a doxycycline-inducible Oct4 transgene have been studied (Hochedlinger et al., 2005). On application of doxycycline, Oct4 is expressed to varying levels in tissues throughout these animals, though not in the brain or testes. On induction of Oct4, mice rapidly developed dysplasia in the stomach, intestines and skin. Notably these tumours appears to derive primarily as a result of inhibition of the differentiation of adult stem cells in these tissues. Remarkably, on withdrawal of doxycycline these tissues rapidly returned to normal, with appropriate differentiation and no sign of teratomas. This suggests that in this context, Oct4 was incapable of reverting these cells to pluripotency. Nonetheless, it would be interesting to investigate whether a

similar block in terminal differentiation and exit from the cell cycle is responsible for the embryonic phenotypes we and Ramos-Mejía *et al.* observed.

It should be noted that in ongoing work I have derived NSC and MEF cell lines from our E12.5 fixed-Oct4 chimeras. This allows us to examine the gene expression profile of these somatic cells and compare them to lines derived from non-chimeric littermates, to generate a greater number of cells to test alternative reprogramming strategies, and to generate a useful resource for future work. In collaboration with Hannah Stuart, I have been working with these cells, and preliminary data suggest that under appropriate conditions it is possible to revert these MEFs to naïve pluripotency solely through manipulating their environment. This builds on our finding that somatic cells from E8.5 chimeras can reprogram. It demonstrates that cells in advanced stages of differentiation can be converted to naïve pluripotency in the presence of an appropriate level of Oct4, and it goes further towards demonstrating that there is not a small, undetected population of pluripotent cells maintained in these embryos as we are able to obtain iPSC lines from a relatively well defined somatic population.

The high contribution of fixed-Oct4 cells to chimeras demonstrates that an ESC level of Oct4 is not incompatible with advanced cellular differentiation, although clearly it interferes with signalling in some contexts. It is interesting to note that in the early embryo, Oct4 is present during each of the major differentiation / cell fate decision events. Maternally contributed Oct4 is present in the zygote, and zygotic Oct4 expression is initiated at around the 4-cell stage. Oct4 can be detected in all cells during the initial segregation of trophectoderm and ICM, though it has been shown that maternal zygotic $Oct4^{-/-}$ embryos successful establish this compartmentalization (Le Bin et al., 2014). Oct4 is then present in all cells of the ICM, and is required for maturation and segregation of the ICM into the naïve epiblast and primitive endoderm. Subsequently, Oct4 is present during the transition and differentiation from the apolar, naïve epiblast to polarized epithelium of the post-implantation epiblast. Studies have shown that Oct4 is required for the equivalent process *in vitro*, where knockout or knockdown of Oct4 prevents this forward progression (Niwa et al., 2000; Radzisheuskaya et al., 2013). Oct4 is still present during the initial stages of gastrulation as cells establish their prospective somatic fate, although the precise role of Oct4 at this point has not been established. Embryos in which Oct4 is deleted around the time of implantation suffer from disorganisation of the epiblast, with poor axis establishment (Mulas et al., 2018), however. Together with the reprogramming data, this suggests

Oct4 plays an important role in facilitating transitions between different cellular identities rather than specifically acting to maintain the pluripotent state.

# Chapter 4: Loss of Oct4 reveals that it is both antagonistic to and necessary for expression of the naïve transcription factor network

## 4.1. Introduction

Oct4 is known to be necessary for the maintenance of pluripotency *in vivo* and *in vitro*. In both cases, loss of Oct4 ultimately leads to exit from pluripotency and eventually cells take on characteristics of the trophoblast lineage, including expression of marker genes Cdx2, Pl-1 and Eomes, and adoption of trophectoderm-like morphology (Nichols et al., 1998; Niwa et al., 2000; Velkey and O'Shea, 2003). Despite this, it is not clear precisely why Oct4 is essential to the naïve state; Oct4 is a highly promiscuous transcription factor, binding to thousands of sites in the genome (King and Klose, 2017; Simandi et al., 2016), and many of its target genes are known not to be essential for the maintenance of ESCs (Hall et al., 2009; Matoba et al., 2006; Zwaka, 2012). At the same time, no target has been identified that can rescue Oct4-loss driven differentiation through overexpression or knockout (Hall et al., 2009; Matoba et al., 2006). It therefore seems likely that differentiation upon loss of Oct4 is the result of misregulation of a number of genes rather than of any single critical target.

Nonetheless, existing models of the action of Oct4 suggest that Oct4 may act on a relatively defined set of targets. The limiting cofactor model suggests that a factor or complex that interacts with Oct4 is present in limiting quantities such that key pluripotency regulatory elements are active normally, but on Oct4 overexpression the cofactor is distributed more broadly resulting in inefficient activation of these important pluripotency genes. Meanwhile the see-saw model proposes that Oct4 is essential for the repression of ectodermal specifiers, counteracting the pro-neural differentiation influence of Sox2.

With this in mind, I set out to investigate early responses to loss of Oct4 in ESCs in order to determine the proximal cause of differentiation. I reasoned that later events would be the cumulative result of a large number of expression changes, while the direct targets of Oct4 responsible for initiating this cascade should be some of the most immediately affected by removing Oct4.

## 4.2. Results

### 4.2.1. Studying Oct4 depletion using a Lox-Cre system

Initially I decided to utilize a conventional conditional Lox-Cre system to study the effect of loss of Oct4 in ESCs. This system uses cell lines containing one floxed Oct4 allele ($Oct4^{F}$) and one knocked out by targeted deletion ($Oct4^{-}$) (Nichols et al., 1998) (Figure 4.1a), which I transfected with a construct for expression of $CreER^{T2}$. In the absence of induction, the oestrogen receptor subunit maintains cytoplasmic localisation of the Cre recombinase, preventing it from interacting with the floxed Oct4 allele in the nucleus. When 4-hydroxytamoxifen (4-OHT) is added and binds to the oestrogen receptor subunit, $CreER^{T2}$ translocates to the nucleus and recombines LoxP sites, resulting in excision of the remaining Oct4 allele ($Oct4^{F} \rightarrow Oct4^{-}$).

Figure 4.1

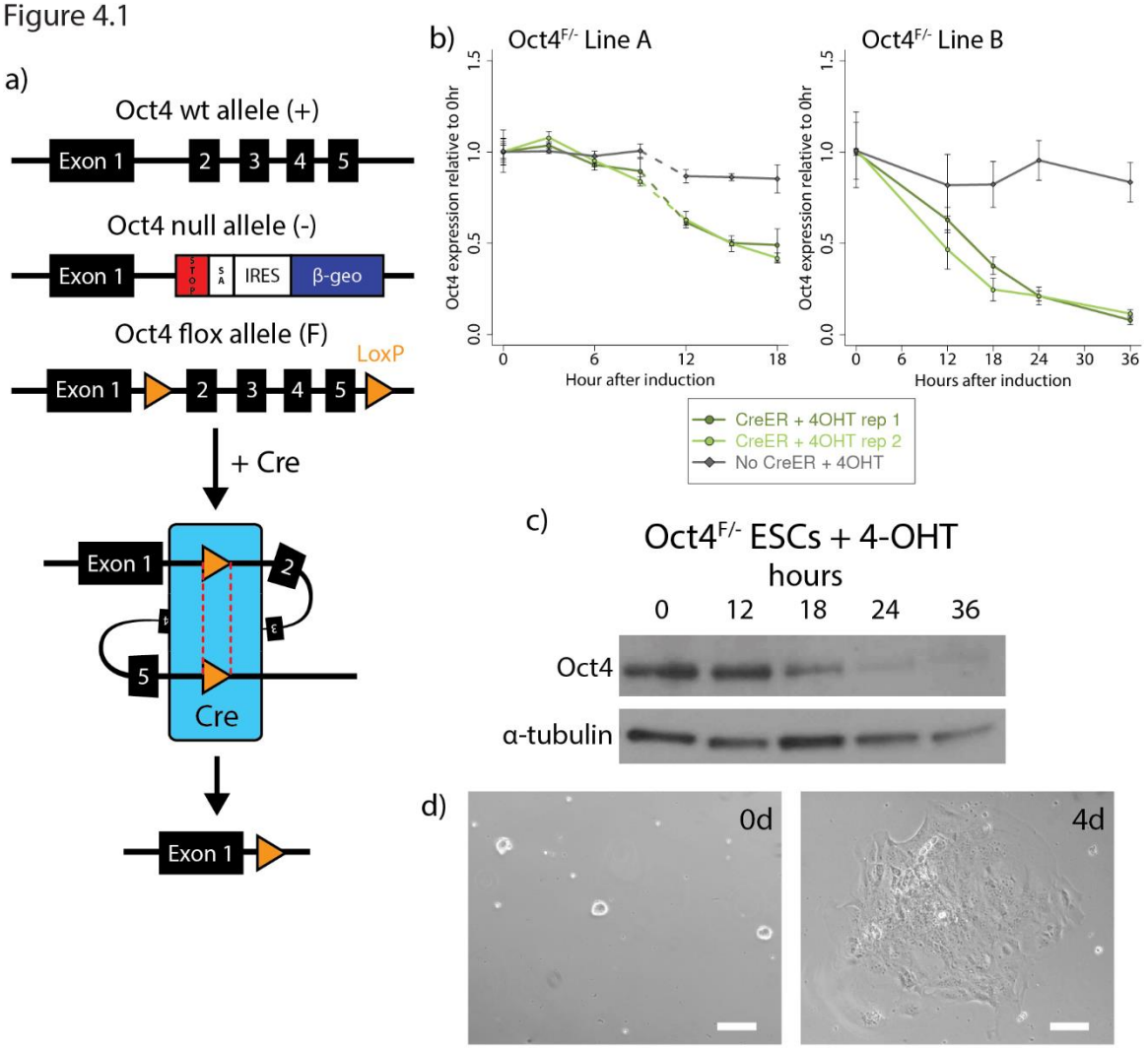



Figure 4.1 Oct4 is depleted from Oct4$^{F/-}$ cells following Cre activation
a) Schematic depicting the wild-type Oct4 allele, the Oct4 null β-geo knock-in allele (Oct4$^{-}$), and the Oct4 floxed exon 2-3 allele (Oct4$^{F}$)which recombines in the presence of Cre to yield an Oct4$^{\Delta Ex2\text{-}5}$ (Oct4$^{-}$) allele. b) Comparison of Oct4 expression, measured by RT-qPCR, in Oct4$^{F/-}$ Line A and Line B following induction of Oct4 deletion by addition of 4-OHT. For each line, parental Oct4$^{F/-}$ cells that had not been transfected with a randomly integrating, constitutively expressed CreERt2 vector were used as a negative control. Two replicate timecourses are shown for each induction. Values represent the mean of three replicate qPCR reactions and error bars show ± SD; values were calculated by the ΔΔCt method using the housekeeping gene Gapdh as a reference gene and normalized to 0 hours for each sample. The dotted lines for Line A represent the fact that later timepoints were induced independently from the earlier timepoints and do not directly follow on in the same induction. Cells were maintained in N2B27-2iL. c) Western blot analysis of Oct4 protein in whole cell lysates from cell line B at time points following addition of 4-OHT. Cells were maintained in N2B27-2iL. d) Phase-contrast images of Oct4$^{F/-}$ ESCs before induction (left) and after 4 days with 4-OHT. Cells were maintained in naïve-specific N2B27-2iL and transitioned to differentiation-permissive SL on addition of 4-OHT to increase long-term survival. Scale bars represent 100μm.

After briefly characterizing cell lines from two genetic backgrounds, I found that one line (Line A, Figure 4.1b) typically only led to Oct4 excision in roughly 50% of cells. Since the second line (Line B, referred to as Oct4$^{F/-}$ ESCs hereafter) gave complete Oct4 depletion, I used this for all further work.

Figure 4.2

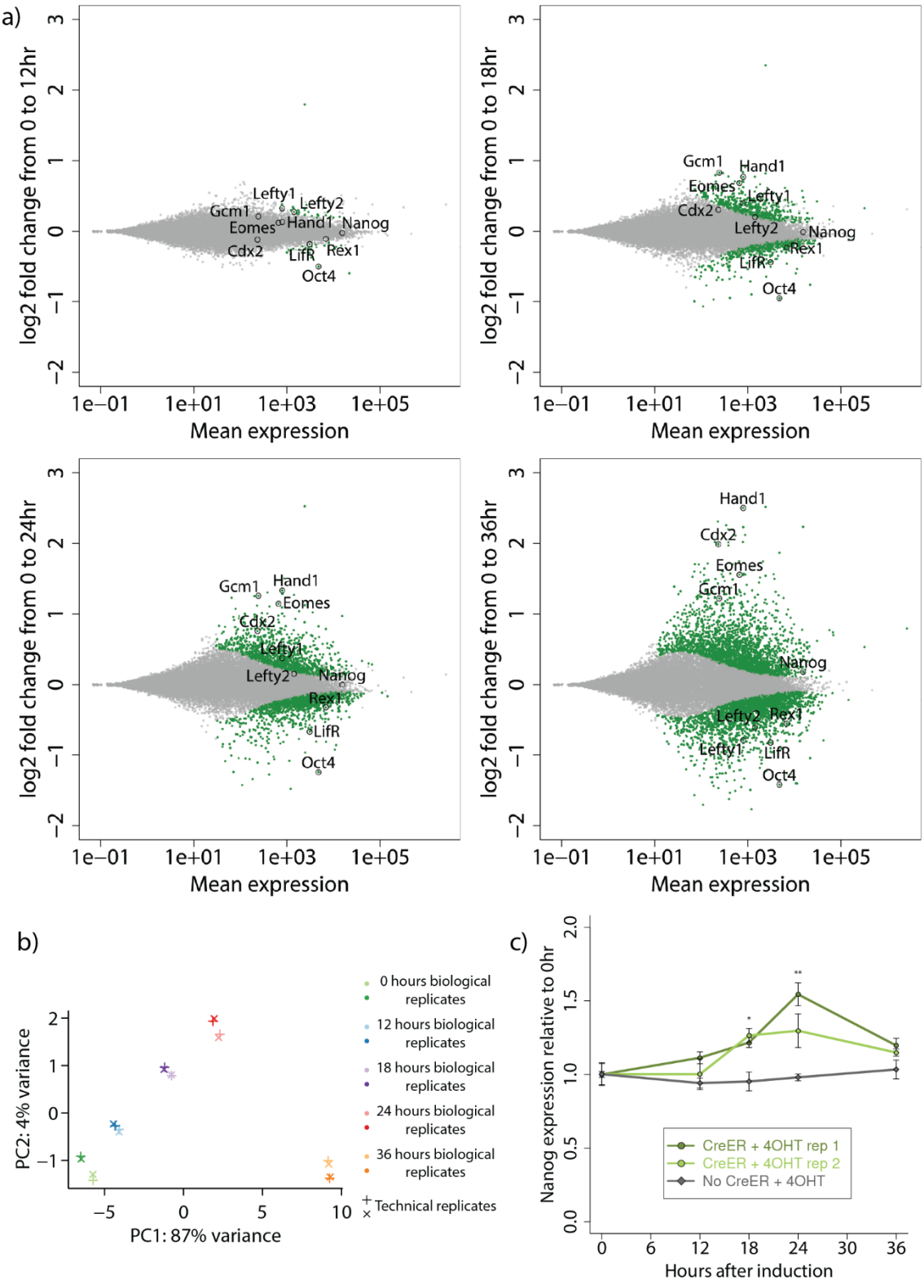



Figure 4.2 Analysis of transcriptional changes following loss of Oct4 in Oct4$^{F/-}$ ESCs

a) Analysis of global gene expression changes following addition of 4-OHT to Oct4$^{F/-}$ ESCs. MA plots showing each gene as a single point. The mean expression of each gene between 0 hours and the indicated timepoint is shown on the x-axis, and the log2 fold change is shown on the y-axis; positive values indicate an increase in expression from 0 hours to the indicated time point and negative values indicate a decrease. Genes showing a statistically significant change (adjusted P-value < 0.05) are highlighted in green. A subset of genes relevant to naïve pluripotency,

trophoblast identity and loss of Oct4 are labelled. Expression normalization and statistical analysis were performed using DESeq2, see Materials and Methods for more details. Number of genes significantly different (adjusted P-value < 0.05) between 0 hours and 12 hours: 55; between 0 hours and 18 hours: 1059; between 0 hours and 24 hours: 2732; between 0 hours and 36 hours: 5233. b) Principal component analysis (PCA) showing the variance between RNA-seq datasets prior to collapsing technical replicated. Note that for each sample there are two biological replicates, and two technical replicates for each of these. PCA performed using DESeq2. c) RT-qPCR analysis of Nanog expression in $Oct4^{F/-}$ ESCs following addition of 4-OHT. For each line, parental $Oct4^{F/-}$ cells that had not been transfected with a randomly integrating, constitutively expressed CreERt2 vector were used as a negative control. Two replicate timecourses are shown for each induction. Values represent the mean of three replicate qPCR reactions and error bars represent ± SD; values were calculated by the ΔΔCt method using the housekeeping gene Gapdh as a reference gene and normalized to 0 hours for each sample. Cells were maintained in N2B27-2iL to minimize perturbation.

Loss of Oct4 protein in these cells is gradual, and is completed within ~36 hours of addition of 4-OHT. Four days after induction, cells began to show morphological changes, resembling trophoblasts then more differentiated trophectodermal cells (Figure 4.1d). I carried out a timecourse over early timepoints after loss of Oct4 in an attempt to identify important direct targets of Oct4. I decided to analyse this data using high throughput RNA-seq methods in order to see global changes in gene expression levels with greater reliability and coverage than traditional microarray approached. Note, principal component analysis of the normalized data prior to collapsing technical replicates shows little variance between technical replicates indicating that there were no problems with the libraries, and little variance between biological replicates indicating the robustness of these changes (Figure 4.2b). Additionally, the principal component of variance maps well with progression in time. As Figure 4.2a shows, in the earliest stages the expression of very few genes changes; within 12 hours of addition of 4-OHT, 55 genes show a statistically significant change in expression. However, few of these are obvious candidates for factors essential to naïve pluripotency. One fifth are general metabolic enzymes with no specific link to pluripotency, which could be responses to the treatment rather than changes in Oct4. A small number of receptors, ligands and transcription factors show expression changes which may imply their regulation by Oct4 in pluripotency. For instance Lefty1 and Lefty2 are upregulated and Fgf4 and Lifr are downregulated on loss of Oct4; direct regulation of the Lefty1 and Fgf4 loci by Oct4 binding have been described before (Dailey et al., 1994; Nakatake et al., 2006; Niwa et al., 2002; Yuan et al., 1995), though Oct4 has previously been reported been reported to promote Lefty1 expression. However, none of these are essential in ESCs. One essential factor downregulated in these cells is G2e3, an E3-ubiquitin ligase. Knockout embryos fail at the blastocyst stage, though this is due to extensive apoptosis rather than trophectodermal differentiation (Brooks et al., 2008).

18 hours after induction, other 1000 genes are differentially expressed. By this point it is impossible to identify any likely key targets over all of the other factors that are changing. That said, it is noteworthy

that several important trophectoderm-associated transcripts are significantly upregulated, including Gcm1, Hand1 and Eomes, while expression of several key pluripotency genes like Nanog remain unchanged (Figure 4.2a). This could suggest that dismantling of the naïve network is a consequence of de-repression of the extraembryonic transcriptional network. By 36 hours, when the majority of Oct4 is finally lost at the protein level (Figure 4.1c), over 5000 genes are differentially expressed. At this time, Lefty1 and Lefty2 are downregulated, as the literature suggests they should be in the absence of Oct4 (Nakatake et al., 2006; Niwa et al., 2002). Surprisingly, however, Nanog expression is increased in these cells (Figure 4.2a), even as they begin to exit pluripotency. This was validated by RT-qPCR, where Nanog expression could be seen to increase as Oct4 protein levels were decreasing (Figure 4.2c).

Direct comparison to the work of Hall et al. (Hall et al., 2009), in which transcriptional repression of Oct4 was used to studying downstream changes, shows similar overall trends. A large amount of variance between the two datasets can be seen in a principal component analysis (Figure 4.3a), likely due to multiple factors such as the use of different cell lines, different growth conditions, RNA treatment and measurement of expression. Despite this, the two systems show very similar trends across the second principal component of variance (Figure 4.3a). Notably, however, progression along the second principal component is more continuous in the Oct4 deletion dataset presented here, whereas the Hall et al. dataset shows a shows less robust separation of earlier timepoints and a large jump to the final timepoint.

As expected, examining the contributions of differentially expressed genes to the principal component analysis shows that timepoints are separated by expression of pluripotency- vs differentiation-associated genes (Figure 4.3b). Comparing the differentially expressed genes in each system, there is a lot of overlap; around a half of the genes that show statistically significant variation in the gene deletion system are also differentially expressed in the transcriptional inhibition system (Figure 4.3c). Oddly, a very large number of genes are differentially expressed even at the earliest timepoint in the Hall et al. dataset, despite there being only a minor change in Oct4 protein level at this time; this agrees with the analysis performed in Hall et al. On the other hand, in the Oct4 deletion dataset there is a gradual accumulation of differentially expressed genes. It is unclear why this is the case, and this is not explained by Hall et al. It may be that there are a number of changes early in the timecourse relating to the induction procedure rather that the downregulation of Oct4, but picking this apart would require a greater amount of data to be generated from control lines by both Hall et al. and myself.

Figure 4.3

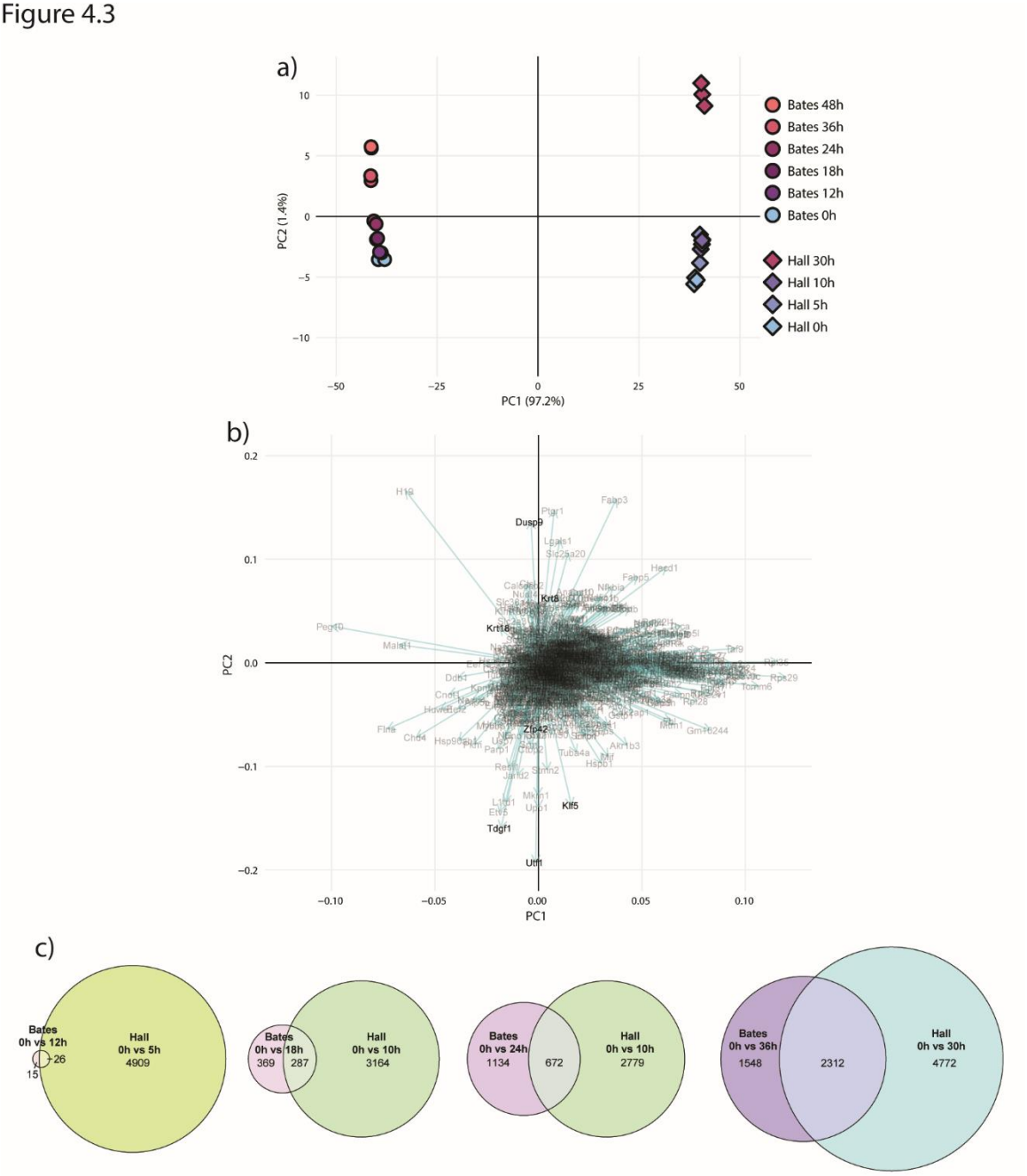



Figure 4.3 Comparison of transcriptional changes following loss of Oct4 in Oct4$^{F/-}$ ESCs vs Oct4 TetOFF ESCs
Comparison of the transcriptional data from Figure 4.2 and that generated by Hall et al. (Hall et al., 2009). a) Principal component analysis of both datasets. Gene annotations for both datasets were converted to Ensembl IDs to allow data to be combined. The top 1000 most variable genes across the complete dataset were normalised with DESeq2 using the variance stabilizing transform, and these values were used to generate the principal components. b) Plot of the loadings for the 1000 genes used to generate the principal components shown in a. Several pluripotency- and differentiation-associated genes are highlighted. c) Euler plots showing the number and overlap of genes showing significantly different expression (adjusted P-value < 0.01) in the two systems at different timepoints.

Despite the similarities between my results and those of Hall et al., these results were somewhat counterintuitive. Some gene expression changes were contrary to published data, and they did not provide any clear clues as to the essential targets of Oct4. However, I noticed that sustained, and even

upregulated, Nanog expression is a feature of Oct4-low cells (Karwacki-Neisius et al., 2013; Radzisheuskaya et al., 2013). It therefore seemed plausible that the unexpected transcriptional changes could be a result of cells transiently passing through an Oct4-low state rather than transitioning directly from a wild-type to a null identity. This is also suggested by the protracted period over which Oct4 protein is lost (Figure 4.1c). The half-life of OCT4 is variously stated at 2-12 hours (Buckley et al., 2012; Lin et al., 2012; Saxe et al., 2009; Wei et al., 2007; Yao et al., 2014), with post-translational modification such as ubiquitilation and sumoylation having a significant impact on the stability of the protein (Wei et al., 2007; Yao et al., 2014). As a result, it is not trivial to study the immediate impact of loss of Oct4 on target genes.

### 4.2.2. Establishing a novel system to study the immediate effects of loss of Oct4

In order to avoid issues with gradual protein decay, I decided to establish a system for depleting Oct4 via conditional proteasomal degradation. The auxin-inducible degron (AID) is a domain found in some auxin-responsive plant proteins (Abel and Theologis, 1996) which, on co-binding the small molecule auxin (IAA), binds the SCF ubiquitin ligase complex (Dharmasiri et al., 2005; Kepinski and Leyser, 2005; Tan et al., 2007). On polyubiquitilation, the AID-containing protein is destroyed by the proteasome. The SCF complex is functionally conserved across eukaryotes (Deshaies, 1999), with expression of different F-box proteins providing spatio-temporal specificity through direct interactions with targets (Skowyra et al., 1997). IAA-bound AID is recognised by the plant-specific F-box protein Tir1 (Gray et al., 1999; Ruegger et al., 1998). Exogenous expression of this protein in mammalian cells allows targeted degradation of any fusion protein containing the AID domain by adding auxin to the culture medium (Holland et al., 2012; Nishimura et al., 2009).

Figure 4.4

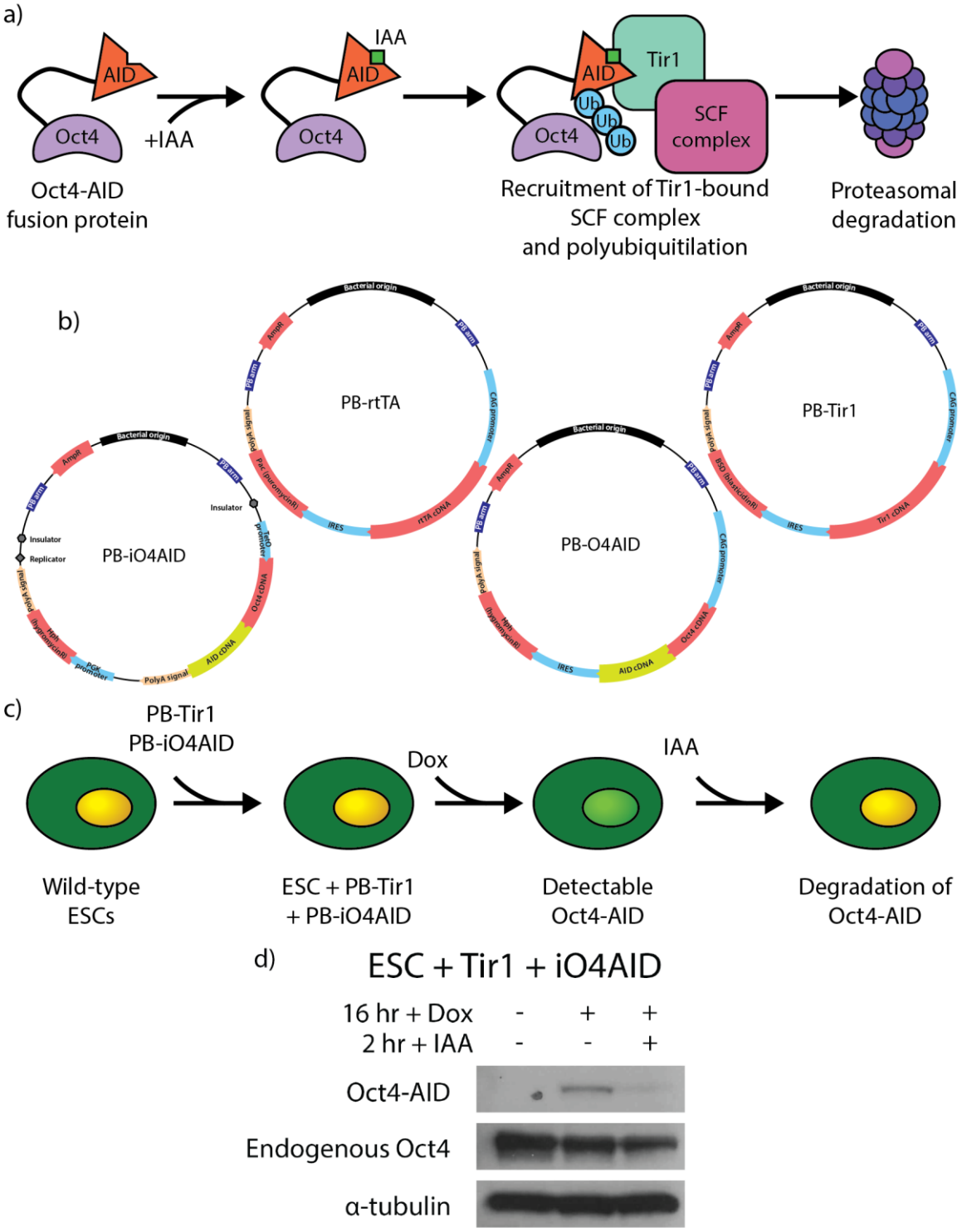



Figure 4.4 Designing and testing a degradable Oct4-AID fusion protein

a) Oct4 is expressed as a fusion protein with the AID domain. When IAA is present, it binds to the AID domain and transgenic Tir1 protein, resulting in polyubiquitilation of Oct4-AID by the endogenous SCF complex. This ubiquitilation results in proteasomal degradation of Oct4-AID. b) Plasmid maps of pPB-CAG-rtTA-IRES-Pac (PB-rtTA), pPB-CAG-NLS-m.OsTir1-IRES-bsd (PB-Tir1), pPB-TRE-O4AID-PGK-Hph (PB-iO4AID), and. c) Schematic for testing degradation of the Oct4-AID fusion protein. Wild-type ESCs were transfected with constructs for Tetracycline response element (TRE) Oct4-AID and doxycycline- (dox-) activated rtTA (together iO4AID) and the F-box protein Tir1. After selection for transgene integration, dox was added to the culture to induce expression of Oct4-AID. While Oct4-AID was induced, IAA was added to cause proteosomal degradation. d) Western blot analysis of Oct4-AID and wild-

type Oct4 in uninduced, dox induced, and dox + IAA induced ESCs previously transfected with iO4AID and Tir1. α-tubulin is shown as a loading control. Dox was added for a total of 16 hours, IAA was added for the final 2 hours. Induction and partial depletion of Oct4-AID fusion protein can be seen.

I designed a construct for the expression of an Oct4-AID C-terminal fusion protein. When co-expressed with Tir1, this should result in rapid depletion on addition of IAA (Figure 4.4a). To test that the AID domain was functional in the fusion protein, I cloned the construct into a Piggybac tetracycline response plasmid to achieve efficient random genomic integration and inducible expression (Figure 4.4b). I transfected this, alongside constitutively expressed Tir1 and rtTA (a transcription factor that activates transcription downstream of a tetracycline response element (TRE) in the presence of doxycycline (dox)) Piggybac vectors, into wild-type mouse ESCs, generating inducible Oct4-AID ESCs (iO4AID ESCs). After selecting for transgene integration, I treated cells with dox to induce expression of Oct4-AID, then added IAA to induce its degradation. As a fusion protein, Oct4-AID has a higher molecular weight than wild-type Oct4, so the two proteins can be readily separated by SDS polyacrylamide gel electrophoresis (SDS-PAGE). Western blot analysis using antibodies specific to the N-terminus of Oct4 shows that the fusion protein is expressed on addition of dox, and largely depleted in the presence of IAA (Figure 4.4c).

Since I had confirmed the degron to be functional, I needed to test that Oct4 was not compromised within the fusion protein. Since Oct4 is essential for the maintenance of pluripotency I decided to confirm this by expressing Oct4-AID in ESCs lacking wild-type Oct4. A TetOFF-Oct4 cell line (ZHBTc4, Oct4-/- TetOFF-Oct4 hereafter) has previously been established through multi-stage genetic manipulation of CGR8 mouse ESCs (Niwa et al., 2000). Both endogenous Oct4 alleles are knocked out in this line. They constitutively express tTA and contain TetOFF controlled Oct4 expression driven from a TRE; on addition of dox, tTA dissociates from the promoter and transcription stops. Functional Oct4 is required for the maintenance of ESCs, and non-functional transgenes cannot rescue self-renewal on dox-induced inhibition of wild-type Oct4 in these cells (Niwa et al., 2002), allowing them to be used in a complementation assay. I transfected these cells with a constitutively expressed Piggybac Oct4-AID and Tir1 constructs to generate Oct4-/- TetOFF-Oct4 Oct4-AID ESCs (Oct4-AID ESCs hereafter) and maintained cells in the presence of dox to inhibit expression of wild-type Oct4 (Figure 4.5a). Cells continued to proliferate and showed normal morphology (Figure 4.5b). Cells expressed the fusion protein with no detectable wild-type Oct4 present (not shown), and did not show substantially altered expression of key pluripotency genes

(Figure 4.5c). This demonstrates that the Oct4 fusion protein retains its original capacity to maintain naïve pluripotency and is not substantially altered in its function by the addition of the AID domain. Finally, I tested that the degron was functional in the context. Where the parental Oct4-/- TetOFF-Oct4 cells downregulated Oct4 after around 3 hours, losing expression 4 hours after induction, OCT4-AID protein was almost undetectable within one hour of addition of IAA (Figure 4.5d).

Figure 4.5

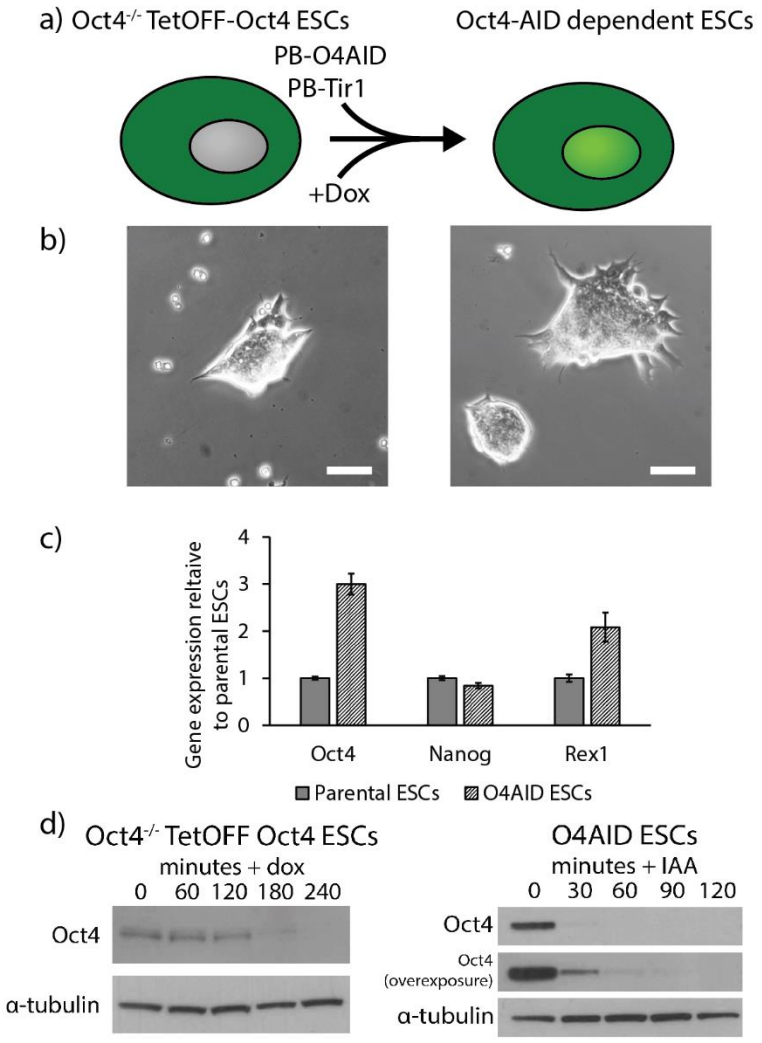



Figure 4.5 Establishing Oct4-AID dependent ESCs (Oct4-AID ESCs)
a) Schematic showing strategy to generate Oct4-AID dependent ESCs. Oct4$^{-/-}$ TetOFF-Oct4 ESCs constitutively express transgenic Oct4, but Oct4 expression is completely abrogated in the presence of dox. These cells were transfected with a randomly integrating piggyBac vector for constitutive expression of the Oct4-AID fusion protein and subsequently maintained in medium containing dox at all times. b) Phase contrast images of Oct4$^{-/-}$ TetOFF-Oct4 ESCs (left) and Oct4-AID ESCs (right). Note that cells maintain a dome-shaped morphology, with similar amounts of spontaneous differentiation between the two lines. Scale bar represents 50μm. Cells were maintained in N2B27-2iL. c) RT-qPCR analysis of expression of Oct4, Nanog and Rex1 in parental Oct4$^{-/-}$ TetOFF-Oct4 ESCs and Oct4-AID ESCs. Values represent the mean of three replicate qPCR reactions and error bars represent ± SD; values were calculated by the ΔΔCt method using the housekeeping gene Gapdh as a reference gene and normalized to parental ESCs. Cells were maintained in N2B27-2iL. d) Western blot analysis of Oct4 levels following depletion from Oct4$^{-/-}$ TetOFF-Oct4 ESCs by addition of dox (left) and Oct4-AID ESCs by addition of IAA (right).

To further validate the function of Oct4 within the fusion protein, I decided to generate iPSCs, null for endogenous Oct4 and therefore wholly reliant on the Oct4-AID transgene for their maintenance. As a starting point, I took the $Oct4^{-/-}$ NSCs used by Aliaksandra Radzisheuskaya to generate fixed-Oct4 cells used in Chapter 3, which had been made by Cre treatment of $Oct4^{F/-}$ NSCs (with the same knockout strategy shown in Figure 4.1a). I transfected these cells with constitutively expressed Piggybac Oct4 or Oct4-AID and transduced them with retroviral Oct4, Klf4 and cMyc. After a period of outgrowth, cells initiated reprogramming and generated intermediate pre-iPSCs (Silva et al., 2008; T. W. Theunissen et al., 2011). I replated these highly proliferative cells in naïve-specific conditions. One of the null Oct4 alleles has a β-geo cassette knocked in which is driven by the endogenous Oct4 regulatory elements. I used this geneticin resistance to select for cells that had activated transcription from this locus in order to ablate non-reprogramming cells. As Figure 4.6b shows, colonies with a domed, iPS-like morphology were readily obtained using either wild-type or Oct4-AID constructs. Pre-iPS reprogramming intermediates showed strong expression of retroviral transgenes, but the pluripotency gene Nanog was not present (Figure 4.6c). In the fully reprogrammed iPSCs, Nanog was robustly expressed while retroviral transgenes were efficiently silenced (Figure 4.6c). The ability of the Oct4-AID transgene to maintain pluripotency in this clean Oct4 knockout background demonstrates that the essential function of Oct4 is maintained. It should also be possible to reprogram cells with Oct4-AID in the absence of any wild-type Oct4 if the fusion construct fully maintains the functionality of the untagged protein; unfortunately this has not been tested thoroughly. In a preliminary experiment in which Oct4-/- NSCs were transfected with wild-type Oct4 or Oct4-AID and transduced with retroviral Klf4 and cMyc, neither were able to generate fully reprogrammed iPS lines. All cells collapsed in 2iL under selection with either form of Oct4. This could indicate that the level of Oct4 used was not appropriate for complete somatic cell reprogramming or that retroviral transduction was inefficient. Since the positive control failed to reprogram in this early work, further experimentation will be required to determine the capacity of Oct4-AID to perform in this role. The generated Oct4-AID iPSCs did not display detectable endogenous Oct4 (Figure 4.6d), while the Oct4-AID fusion protein was robustly expressed as in the Oct4-AID ESCs. Again, I checked that the AID domain was functional, and indeed Oct4 was almost undetectable within 2 hours of IAA treatment.

Figure 4.6

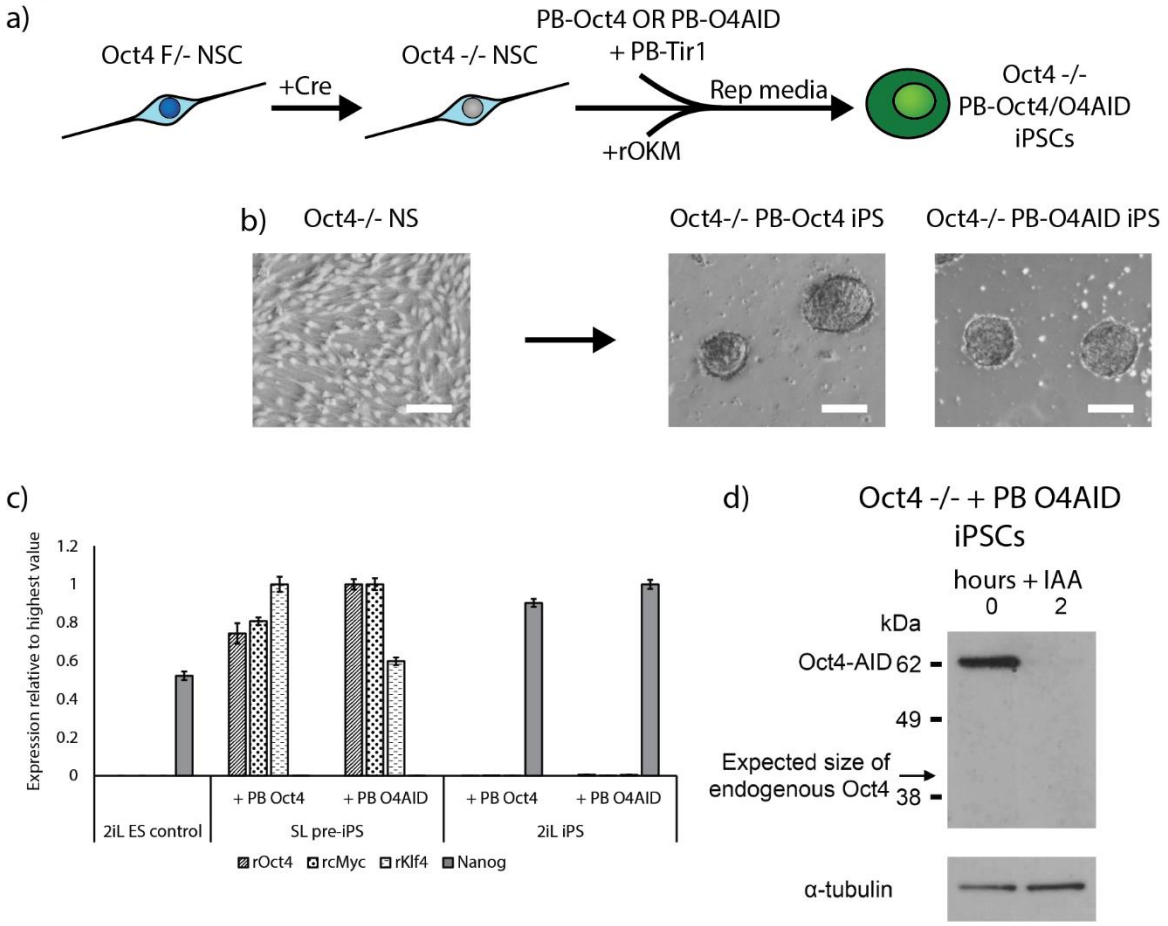



Figure 4.6 Establishing Oct4-AID dependent iPSCs (Oct4-AID iPSCs)
a) Schematic of the approach taken to generate iPSCs dependent on the Oct4-AID construct. Previously, $Oct4^{F/-}$ NSCs were derived and the remaining Oct4 allele was removed via Cre treatment by Aliaksandra Radzisheuskaya. Cells were transfected with a randomly integrating piggyBac vector containing constitutively expressed wild-type Oct4 or the Oct4-AID fusion protein. Subsequently cells were transduced with retroviruses encoding Oct4, Klf4 and cMyc transgenes. Cells were transitioned to SL conditions for several days until pre-iPSC reprogramming intermediates emerged. Cells were subsequently passaged into reprogramming media (KSR-2iL) until colonies with an iPSC morphology were identifiable, at which point selection for activation of the endogenous Oct4 locus was applied. Once stable iPSC lines were established, cells were transitioned to naïve-specific N2B27-2iL. b) Initial NSCs (left) successfully reprogrammed with either Oct4 or Oct4-AID constructs, yielding iPSC lines with normal morphology (right) that could be serially passaged and survived selection for active transcription from the endogenous Oct4 locus. Scale bars represent 100μm. c) RT-qPCR analysis demonstrating complete silencing of retroviral transgenes and expression of the pluripotency marker Nanog at levels equivalent to wild-type ESCs. Taqman assays specific to the retroviral Oct4, Klf4 and cMyc transgenes (rOct4, rKlf4, rcMyc) were used to distinguish retroviral expression from endogenous expression. Pre-iPSC intermediates robustly express retroviral transgenes but have not yet activated the naïve transcriptional network. Transition to mature iPSC state is marked by silencing of retroviral transgenes and expression of markers of naïve pluripotency. Values represent the mean of three replicate qPCR reactions and error bars represent ± SD; values were calculated by the ΔΔCt method using the housekeeping gene Gapdh as a reference gene and normalized to the sample with greatest expression of each gene. d) Western blot analysis of Oct4-AID iPSCs showing expression of the fusion protein independent of the wild-type protein, and efficient depletion on addition of IAA.

### 4.2.3. Using rapid depletion to probe the essential function of Oct4 in ESCs

With this system that allows rapid depletion of Oct4 at the protein level, I planned to identify the essential role of Oct4 in ESCs. While existing systems did not show full depletion of Oct4 until long after transcriptional changes were already evident, the Oct4-AID system achieves complete loss of Oct4 in such a short space of time that I expected cells to avoid passing through the confounding Oct4-low state. I therefore hoped that this system would allow me to identify early transcriptional changes caused by Oct4 deletion that correspond to essential targets of Oct4 in the maintenance of naïve pluripotency.

I performed qRT-PCR analysis of expression of several pluripotency genes in a timecourse following acute depletion of Oct4 in these ES and iPSCs. Many naïve-associated factors were rapidly downregulated on loss of Oct4 (Figure 4.7a-b). While some of these downregulated genes, such as Nr0b1, were repressed relatively rapidly in $Oct4^{F/-}$ ESCs on genomic ablation of Oct4, others such as Nanog appeared to be actively maintained in the slower system. This shows that there are important differences between reducing Oct4 levels and removing Oct4 entirely in terms of transcriptional responses. Examining several trophoblast-associated genes confirmed that these cells exit pluripotency following loss of Oct4 and differentiate towards trophectodermal lineages (Figure 4.7a-b), in keeping with conventional Oct4 depletion systems.

Figure 4.7

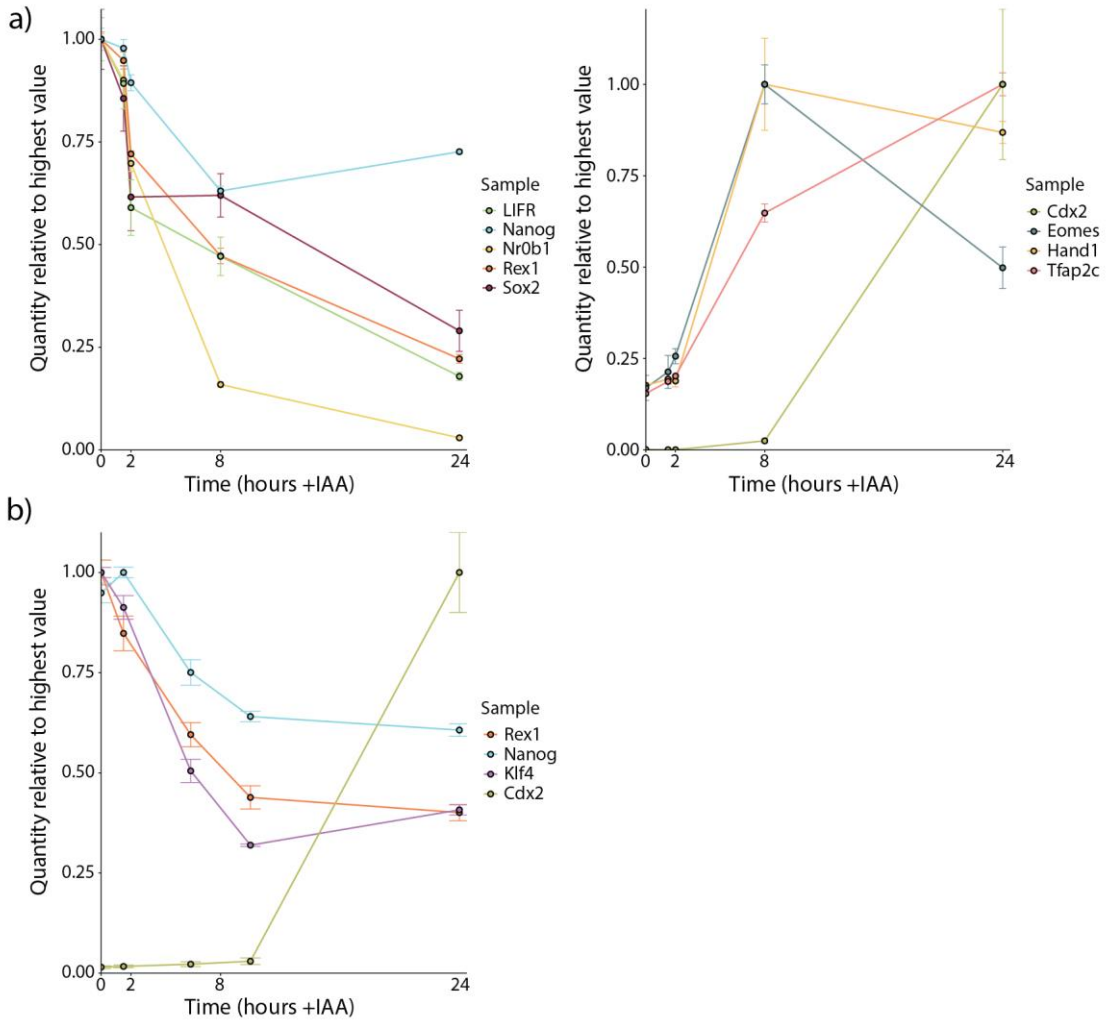



Figure 4.7 Profiling transcriptional changes in Oct4-AID ESCs and iPSCs following depletion of Oct4
a - b) RT-qPCR analysis of several pluripotency factors immediately following Oct4 depletion by addition of IAA in a) Oct4-AID ESCs and b) Oct4-AID iPSCs. Values represent the mean of three replicate qPCR reactions and error bars represent ± SD; values were calculated by the ΔΔCt method using the housekeeping gene Gapdh as a reference gene and normalized to 0 hours. Cells were maintained and induced in naïve-specific N2B27-2iL conditions to minimize perturbations.

Global transcriptional changes were examined by RNA-sequencing, though only a single sample for each timepoint was analysed. Consequently, statistical and advanced bioinformatics analysis is impossible. A conservative analysis suggests that cells do not undergo conventional exit from pluripotency on loss of Oct4. In typical differentiation from naïve pluripotency cells begin to lose naïve-specific gene expression, but maintain expression of general pluripotency genes such as Sox2 and Tdgf1 (Brons et al., 2007; Kalkan et al., 2017; Tesar et al., 2007). Coincident with this is upregulation of early post-implantation genes, which are upregulated in primed pluripotent EpiSCs relative to naïve ESCs (Brons et al., 2007; Hayashi et al., 2011; Kalkan et al., 2017; Tesar et al., 2007). Subsequently, lineage markers are upregulated (Han et al., 2010; Joo et al., 2014; Kojima et al., 2014), prior to loss of pluripotency gene expression and lineage commitment. However, in the case of

differentiation induced by loss of Oct4, we observe concurrent loss of both naïve-specific and general pluripotency gene expression (Figure 4.8a). There is no notable increase in expression of markers of primed pluripotency or of the early post-implantation epiblast. Less surprisingly, subsequent upregulation of embryonic lineage markers never occurs either. Instead, we observe upregulation of extraembryonic genes, particularly trophoblast-associated transcripts (Figure 4.8a), though the earliest changes remain the rapid downregulation of specific naïve transcription factors. Comparing global transcriptional changes between the genetic ablation system and induced protein degradation system, it is clear that cells follow the same trajectory of transcriptional change, indicating that the overall pattern of differentiation is comparable (Figure 4.8b). However, the AID system yields far more rapid changes.

Figure 4.8

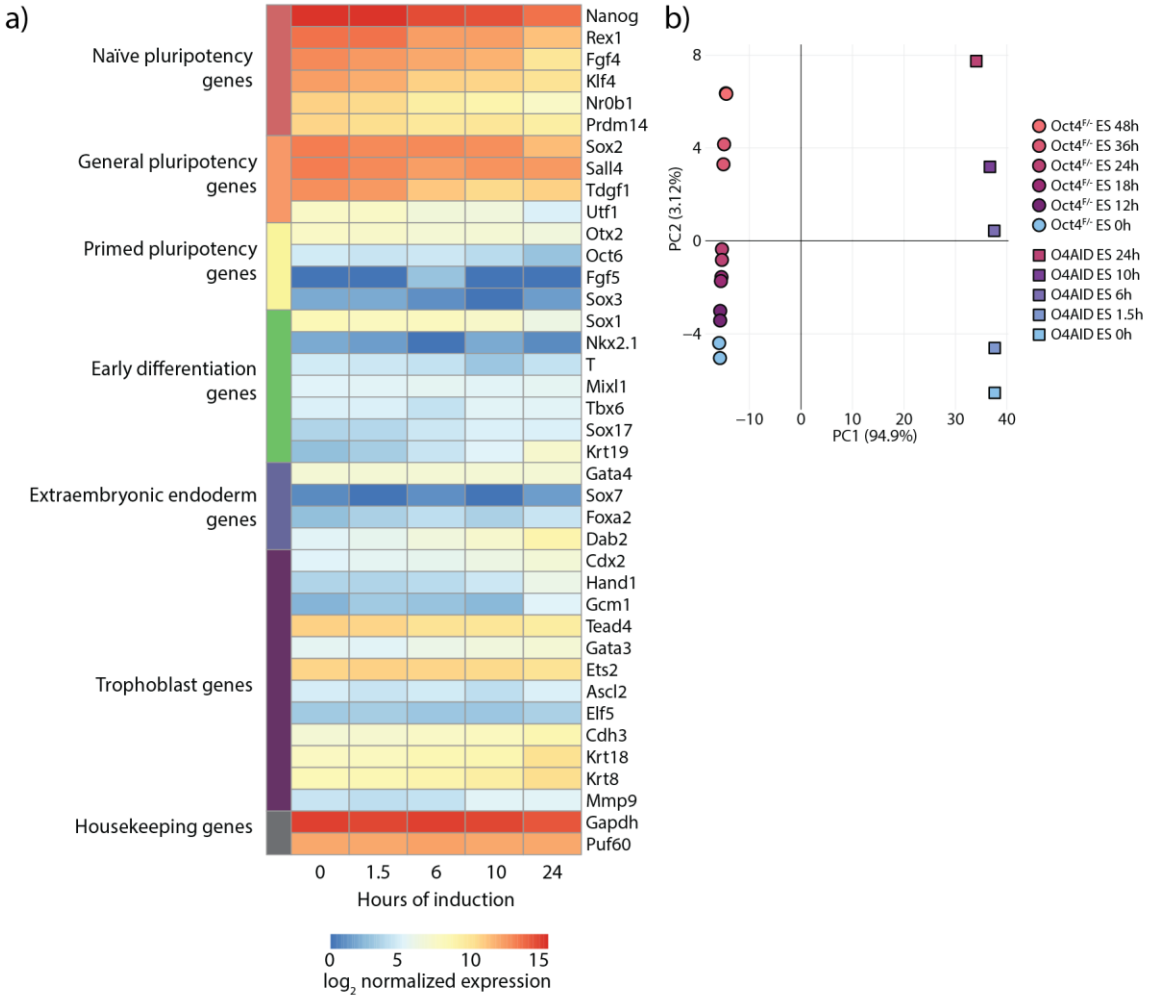



Figure 4.8 Global transcriptional changes in Oct4-AID ESCs reveals non-conventional exit from pluripotency
a) Heatmap showing expression of genes representative of several cell identities in Oct4-AID ESCs over a timecourse following addition of IAA. Intensities represent $log_2$ normalized expression calculated from RNA-seq data using DESeq2. There is a single replicate per timepoint. Note that unlike descriptions of conventional exit from pluripotency,

there is no increase in expression of primed pluripotency genes coincident with the decrease in expression of naïve pluripotency expression. Instead, expression of a subset of trophoblast markers genes are upregulated. b) PCA plot showing variance among and between samples from $Oct4^{F/-}$ ESCs in Figure 4.2 and Oct4-AID ESCs following Oct4 deletion.

### 4.2.4. Studying the role of Oct4 in the binding of other transcription factors to DNA

Many pluripotency-associated transcription factors bind to a large number of enhancer elements throughout the genome (Kim et al., 2008; Loh et al., 2006), including many genes that are not actively transcribed in ESCs. To explain this phenomenon, it has been suggested that binding of several transcription factors may be required for the activation of genes in naïve cells. This is supported by the fact that many active loci in ESCs display binding of multiple pluripotency factors at their enhancer elements (Chen et al., 2008; Göke et al., 2011). With this in mind, it has been hypothesised that a key role of Oct4 may be to recruit other transcription factors to pluripotency genes (King and Klose, 2017), either through direct interactions or by creating a less restrictive chromatin environment through epigenetic modifications.

It has been found that Oct4 and Sox2 form a complex on DNA (Fong et al., 2011; Kuroda et al., 2005; Lam et al., 2012; Merino et al., 2014), and bind to their respective motifs in a cooperative manner (Ambrosetti et al., 2000, 1997; Reményi et al., 2003; Tokuzawa et al., 2003). Oct4 has also been shown to form complexes with Nanog (Liang et al., 2008; Pardo et al., 2010; Wang et al., 2006), and they co-bind to many loci throughout the genome (Chen et al., 2008; Göke et al., 2011; Kim et al., 2008; Loh et al., 2006). I was therefore interested to see how the binding of these pluripotency transcription factors to the genome would be affected immediately following loss of Oct4.

Western blot analysis confirmed that 90 minutes after addition of IAA, Oct4 was efficiently depleted while the level of Sox2 and Nanog proteins were yet to be strongly affected in Oct4-AID ESCs (Figure 4.9a). Pulling down regions associated with Sox2 by chromatin immunoprecipitation (ChIP) and analysing enrichment by qPCR revealed a small decrease in Sox2 binding to the Oct4 distal enhancer (Figure 4.9b). This did not achieve statistical significance due to the small number of replicates, so it should be interpreted cautiously, but is in agreement with previous studies (Ambrosetti et al., 2000, 1997; King and Klose, 2017; Reményi et al., 2003; Tokuzawa et al., 2003). On the other hand, the immediate response of Nanog to loss of Oct4 was an increase in binding to the Oct4 regulatory region, though again this data requires replication to give confidence to these findings (Figure 4.9c). A previous study using TetOFF-Oct4 ESCs suggested that Nanog binding decreased in the absence of Oct4 (King and Klose, 2017); however, in that work ChIP was performed 24 hours after induction of

Oct4 depletion, leaving ample time for secondary effects to alter the Nanog binding profile. Interestingly, other work has suggested that in the Oct4-low state Nanog has increased genomic occupancy, though it is unclear whether this is associated with increased or decreased Oct4 at the same sites (Karwacki-Neisius et al., 2013; Radzisheuskaya et al., 2013). In the data presented here, we actively avoid the Oct4-low state; this common finding suggests that Oct4 may indeed actively reduce the ability of Nanog to bind to enhancer elements.

Figure 4.9

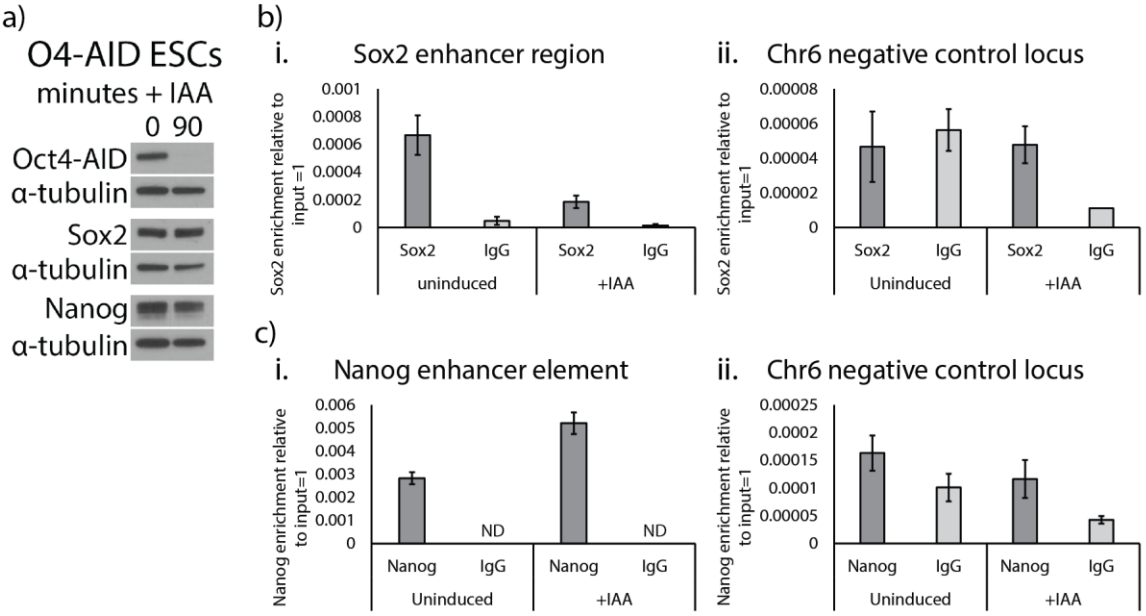



Figure 4.9 Oct4 affects the binding of other pluripotency factors to the genome
a) Western blot analysis of Oct4-AID fusion protein, Sox2 and Nanog in Oct4-AID ESCs before and two hours after addition of IAA. b) ChIP-qPCR analysis of Sox2 binding to i. a proximal downstream enhancer element of Sox2 and ii. a gene desert known to lack Oct4, Sox2 or Nanog binding sites on chromosome 6 before and two hours after depletion of Oct4. c) ChIP-qPCR analysis of Nanog binding to i. the Nanog enhancer element and ii. the gene desert. Values represent the mean of three IPs for Sox2/Nanog and the mean of three replicate qPCR reactions for IgG, and error bars represent ± SD; values were calculated by the ΔCt method using an input control for each sample such that 100% of input = 1. Cells were maintained and induced in N2B27-2iL with dox.

To further validate this, I repeated the experiment in Oct4-AID iPSCs. Again, addition of IAA rapidly and efficiently depleted Oct4 from cells with minimal impact of Nanog protein levels (Figure 4.10a). In this experiment there was little change in Nanog binding to the Oct4 distal regulatory element, but there was a significant increase in the pulldown efficiency of the Nanog enhancer (Figure 4.10b). This confirmation of increased Nanog bound to DNA in an independent cell line lends further strength to this dataset.

Figure 4.10

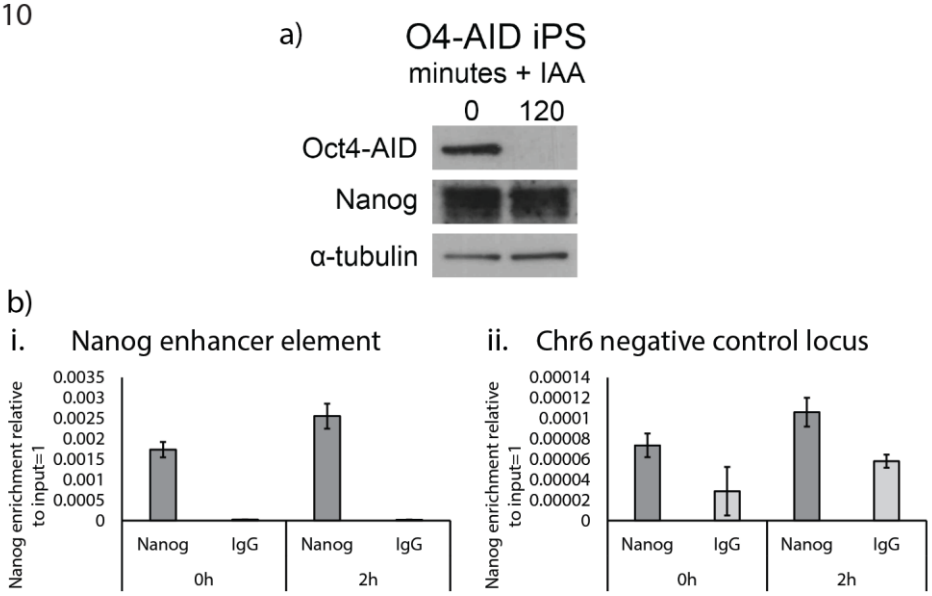



Figure 4.10 Nanog binding increases at some loci following Oct4-AID depletion in Oct4-AID iPSCs
a) Western blot analysis of Oct4-AID and Nanog in Oct4-AID iPSCs before and two hours after addition of IAA. b) ChIP-qPCR analysis of Nanog binding to i. the Nanog enhancer element and ii. a gene desert known to lack Oct4 or Nanog binding sites on chromosome 6, before and two hours after depletion of Oct4. Values represent the mean of three IPs for Nanog and the mean of three replicate qPCR reactions for IgG, and error bars represent ± SD; values were calculated by the ΔCt method using an input control for each sample such that 100% of input = 1. Oct4-AID iPSCs were maintained and induced in N2B27-2iL.

### 4.2.5. Global analysis of Nanog distribution following loss of Oct4

To test the impact of loss of Oct4 on Nanog localisation globally, I submitted Nanog immunoprecipitated chromatin for high thoughput sequencing. After aligning the reads to the mouse genome, I examined the Nanog binding profile at the regulatory regions of several key pluripotency genes using the UCSC Genome Browser (Kent et al., 2002). There was a clear increase in the read density, represented by a greater peak area, following depletion of Oct4 (Figure 4.11). This confirms that Nanog binding increases at key pluripotency loci in the absence of Oct4.

Figure 4.11

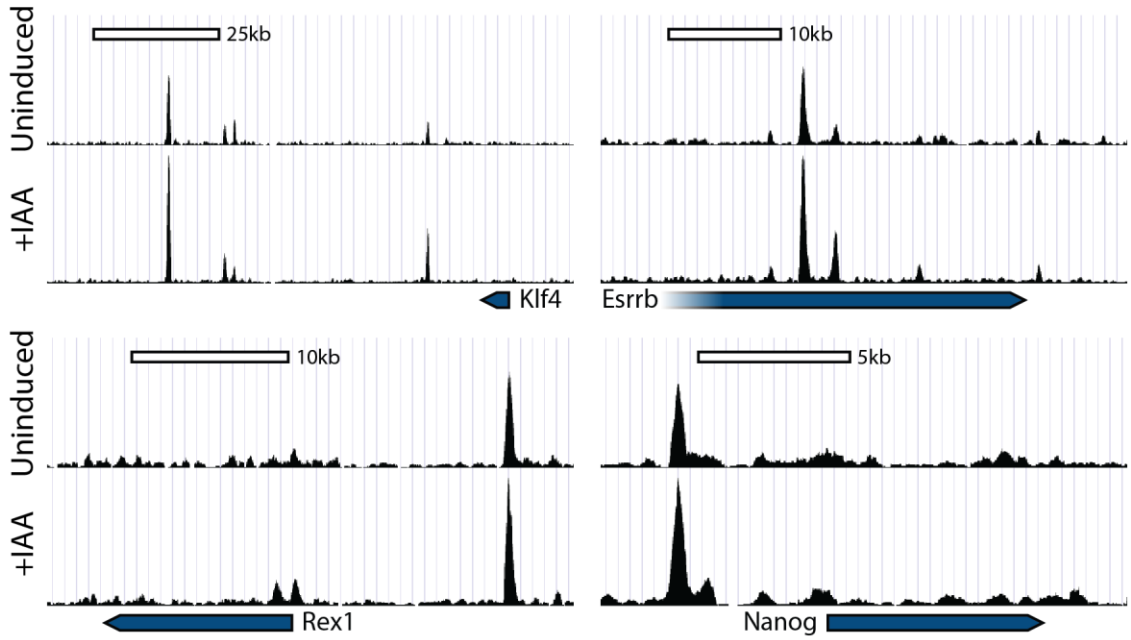



Figure 4.11 Nanog genomic occupancy at key pluripotency regulatory elements
Nanog ChIP-seq in Oct4-AID ESCs displayed on the UCSC genome browser in uninduced and IAA-induced Oct4-AID ESCs at several representative genomic loci. Each track was generated by merging the sequencing output of 2 IPs. The y-axis represents read counts normalized to library size, and scaled to the maximum peak height at each locus using the same scale for each sample. Peaks can be seen at enhancer elements, and peak height is greater in Oct4-depleted samples. ChIP was performed using a Nanog-specific antibody. Using the output DNA, NGS library preparation was performed by the SCI genomics facility, and sequencing was performed at the CRUK CIGC. I performed subsequent analysis. Oct4-AID ESCs were maintained and induced in N2B27-2iL conditions.

I decided to look globally at all Nanog peaks, generated from the Oct4-depleted samples since they appeared to show increased Nanog binding. By plotting the change in read counts at all 12622 loci, it is clear that Nanog binding increases at the majority of binding sites, though it does decrease at a small number of loci (Figure 4.12a). I also looked more specifically at the complementary subsets of peaks that either overlap with an Oct4 peak (1812 sites) or do not co-bind with Oct4 (10810 sites) (Figure 4.12a-b). At Oct4-Nanog co-binding sites there was a detectable increase in overall Nanog binding; globally this was statistically significant (Mann-Whitney test, $P < 10^{-10}$), and many individual peaks did show greater binding (52 increased and 6 decreased with an FDR <0.05 out of a total of 1810 peaks). However, Nanog peaks that do not overlap with Oct4 displayed a more dramatic increase in Nanog binding (Mann-Whitney test, $P < 10^{-10}$), with more than 10% of peaks showing greater Nanog occupancy in the absence of Oct4 (1650 increased and 2 decreased with an FDR <0.05 out of a total of 10753 peaks). Examining heatmaps of Nanog enrichment (Figure 4.12bi. and ii.), it appears that Nanog is enriched at all of the binding sites identified in Oct4-depleted samples prior to removal of Oct4. This

suggests that Nanog is not binding to new loci in the absence of Oct4, even if the overall level of binding is increasing.

While counterintuitive, this implies that Oct4 may be able to regulate the ability of Nanog to interact with chromatin indirectly, perhaps through sequestering a binding partner or Nanog itself in the nucleoplasm. Alternatively, it may be that Nanog is recruited to sites that it only weakly binds to, and are therefore undetected, by Oct4, and thus in the absence of Oct4 this pool of Nanog would be free to associate with stronger binding sites; however, it seems unlikely that this is sufficient to account for the level of increased binding observed. It would be interesting to more precisely quantitate the amount of Nanog bound to chromatin, either by relative quantitation in chromatin-enriched samples by mass spectrometry or through single-molecule imaging of Nanog and determination of the chromatin-bound fraction by their residence time, in order to clarify whether there is a broad increase in genomic occupancy by Nanog, or whether the changes observed represent a redistribution across Nanog binding sites.

Figure 4.12

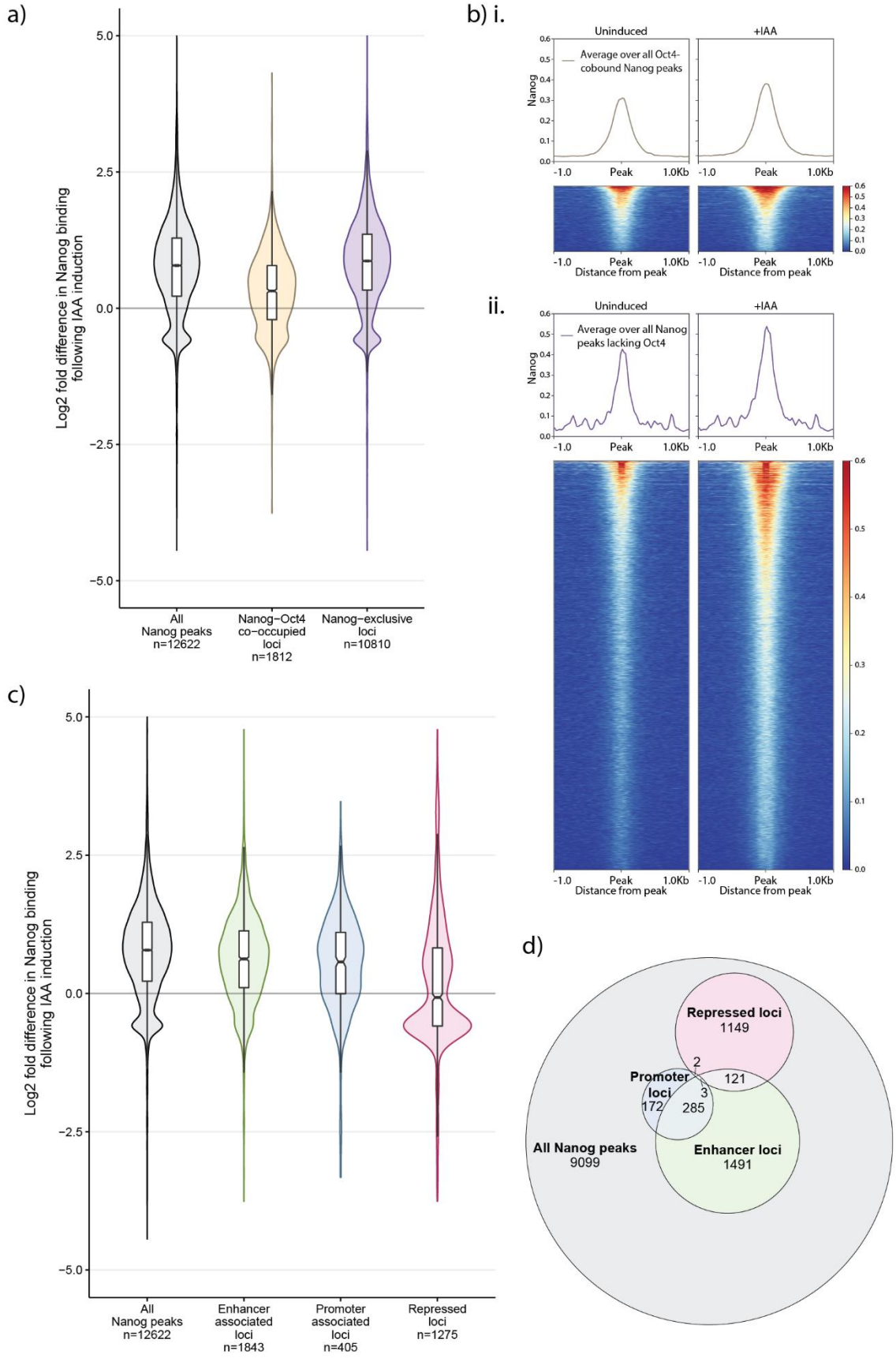



Figure 4.12 Global analysis of Nanog binding sites

Analysis of Nanog occupancy globally and at Nanog binding sites that are cobound by Oct4 or do not show enrichment for Oct4 (right). a) Violin and box plots showing the log2 fold change in normalised read counts in induced vs uninduced O4-AID ESCs at every Nanog peak (n=12622), and split into those that are co-bound by Oct4 (n=1812) or lack Oct4 (n=10810). b) Top: summary plots showing the mean Nanog binding intensity (counts per million) 1kb either side of Nanog binding sites that normally i. do or ii. do not display co-bound Oct4 in uninduced and Oct4-depleted Oct-AID ESCs. Bottom: heatmaps showing Nanog binding intensity 1kb either side of Nanog binding sites that normally i. do or ii. do not display Oct4 co-binding. Intensities correspond to values shown to the side in counts per million. c) Violin and box plots showing the log2 fold change in normalised read counts in induced vs uninduced O4-AID ESCs, shown for different genomic contexts. Log-fold-change is shown for all Nanog peaks (n=12622), those at loci enriched for signs of active enhancer elements (1843), those at loci enriched for promoter associated marks (405), or those at loci enriched for repressive marks (1275). d) Euler plot showing the overlap of Nanog peaks with marks of active enhancers, active promoters, and repressed loci. Nanog binding sites were calculated using MACS2 with reads from the +IAA dataset as this had a greater signal to noise ratio. Oct4 binding sites were calculated against whole-cell extract background using MACS2 with publically available datasets (Marson et al., 2008). Nanog binding sites at active enhancers were defined as loci that both overlap by at least 1bp with H3K27ac and are within 1kb of a p300 ES binding site; Nanog binding sites at active promoters were defined as loci that both overlap by at least 1bp with H3K4me3 and are within 1kb of a PolII binding site; Nanog binding sites at repressed loci were defined as loci that overlap with either H3K9me3 or H3K27me3. Publically available peak-called ChIP data for H3K4me3, H3K9me3, H3K27ac, H3K27me3, p300 and PolII in Bruce4 mouse ESCs were downloaded from the ensembl website. Oct4-AID ESCs were maintained and induced in N2B27-2iL with dox.

I additionally examined the effect of chromatin context on the change in Nanog binding following loss of Oct4. Using publically available data for the binding of histone marks associated with active enhancer or promoters or repressed loci, as well as the highly abundant transcription factor p300 and RNA PolII, I assigned Nanog binding sites as associated with one or more of these loci. As expected, the largest group was enhancer associated Nanog binding sites (Figure 4.12d). Looking at the change in Nanog binding following loss of Oct4, there is an increase in binding at many enhancer- and promoter-associated sites (Figure 4.12c). However, at repressed loci there appears to be a bimodal distribution of changes in Nanog binding; a small number of sites show very large increases in Nanog binding, but loci in which Nanog binding decreases appear to be enriched in this group (Figure 4.12c).

One possible explanation for an increase in enhancer-bound Nanog would be if the protein is more stable in the absence of Oct4. To test whether this is the case I performed a cycloheximide chase assay (Figure 4.13a). Cycloheximide inhibits protein synthesis by binding within ribosomes and preventing translocation by blocking the passage of tRNAs out of the complex (Schneider-Poetsch et al., 2010). As a result, no new proteins are synthesized, and by monitoring the levels of proteins over time their rate of decay can be determined. I quantitated Nanog levels following addition of cycloheximide, with or without addition of IAA, by Western Blot, using a standard curve to adjust for the non-linear relation between protein quantity and band intensity (Figure 4.13b). As expected, Nanog degradation followed exponential decay kinetics and as a result can be described by Equation 1. Plotting the natural log of the protein quantity yields the linear equation in Equation 2, so the decay constant, $\lambda$, can be

calculated as the gradient of the line. Doing this gives the decay constants shown in Figure 4.13b-iii. Finally, using Equation 3, we can calculate the half-life of Nanog; in the presence of Oct4 this is 2.3 hours, and this is reduced to 1.4 hours in the absence of Oct4. This fits with previous findings that Nanog displays increased stability in the presence of Oct4 (Muñoz Descalzo et al., 2013). Again, however, this previous study investigated Nanog stability 24 hours after the initiation of Oct4 depletion, whereas the work presented here examines the immediate impact of loss of Oct4. Regardless, this suggests that stability is not the cause of the increased genomic occupancy of Nanog that we observed in Figures 4.9-4.12.

Figure 4.13

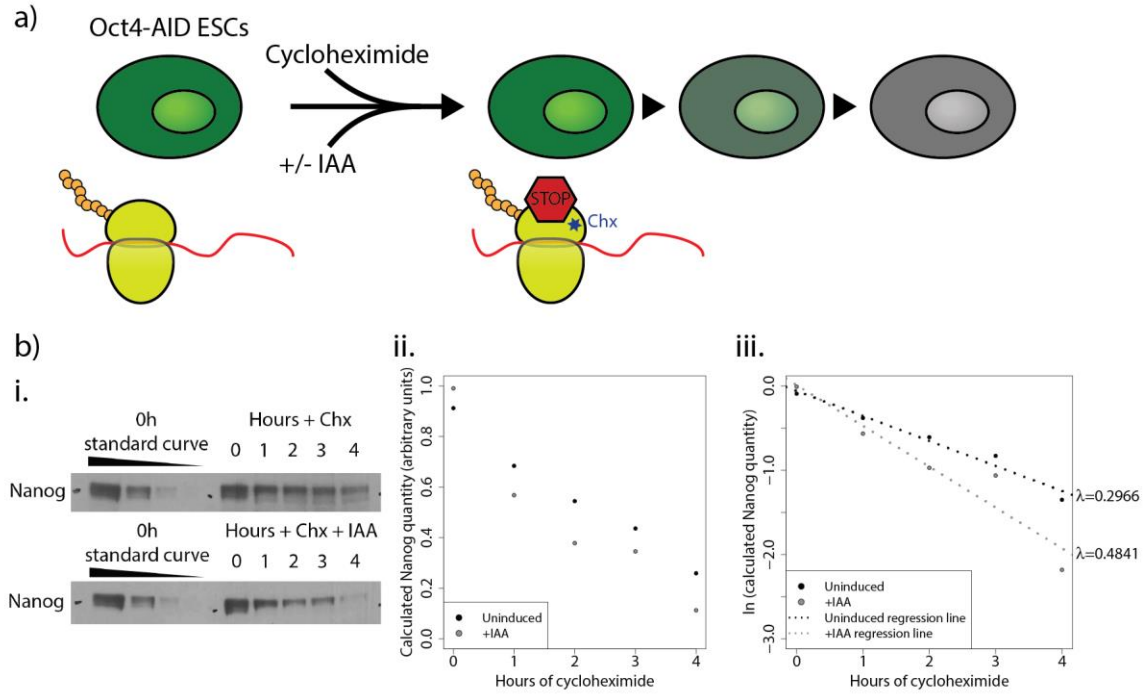



Figure 4.13 Nanog protein is more stable in the presence of Oct4-AID
a) Schematic of the experiment. Oct4-AID ESCs were plated in replicate and cycloheximide + or - IAA was added. Samples were taken every hour for 4 hours. Since cycloheximide blocks translation by ribosomes, protein levels decrease over time as they are degraded. b) Analysis of Nanog protein half-life in presence and absence of Oct4. i. Western blot of Nanog protein used for further analysis. A standard curve was used to calculate protein quantity from band intensity using ImageJ 1.52e. ii. Graph of calculated Nanog quantity over time following addition of cycloheximide + or - IAA. The data show exponential decay kinetics. iii. Graph showing the natural log of Nanog quantity over time. A straight line was fit to the data using regression analysis (Uninduced $R^2$=0.9683; +IAA $R^2$=0.9148) with gradients equal to the decay constant indicated to the right. Oct4-AID ESCs were maintained and induced in SL to minimize cell death.

$$x(t) = x_0 e^{\lambda t}$$

Equation 1: Exponential decay equation
The quantity of a protein undergoing exponential decay can be described as a function of time, x(t), where $x_0$ is the initial concentration of the protein and λ is the decay constant.

$$\ln\big(x(t)\big) = \ln(x_0) + \lambda t$$

Equation 2: Natural log of Equation 1
Taking the natural log of Equation 1 creates a linear equation, making it trivial to calculate the decay constant, λ, as it is the gradient of a straight line.

$$t_{1/2} = \frac{\ln(2)}{\lambda}$$

Equation 3: Relationship of half-life and decay constant
The half-life of a protein, $t_{1/2}$, is equal to the natural log of 2 divided by the decay constant, λ.

### 4.2.6. Studying the epigenetic impact of loss of Oct4

Given this increase in Nanog binding to pluripotency genes, including at enhancer elements where it acts as a transcriptional activator (Festuccia et al., 2012; Shi et al., 2006; Stuart et al., 2014), I was interested to see if there were epigenetic changes associated with increased enhancer usage in cells following loss of Oct4.

Figure 4.14

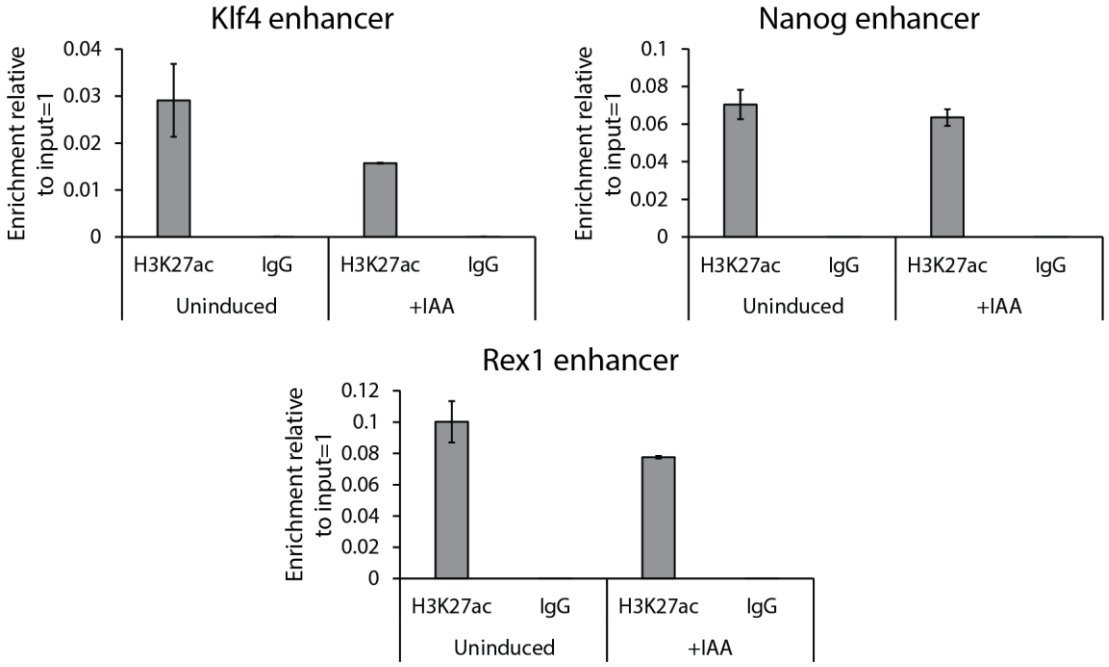



Figure 4.14 Changes in active histone modification H3K27ac following Oct4 depletion in Oct4-AID ESCs
ChIP-qPCR analysis of H3K27ac enrichment at the regulatory elements of Klf4, Nanog and Rex1 before and two hours after depletion of Oct4 by addition of IAA. Values represent the mean of two IPs and error bars represent ± SD; values were calculated by the ΔCt method using an input control for each sample such that 100% of input = 1. Oct4-AID ESCs were maintained and induced in SL + dox.

Acetylation of histone H3 at lysine 27 (H3K27ac) is a marker of active enhancer elements; surprisingly, despite the increase in Nanog binding, H3K27ac enrichment did not increase at the regulatory elements of several pluripotency genes short after depletion of Oct4, instead decreasing at several loci (Figure 4.14).

Given this apparent decrease in enhancer usage, despite increased Nanog occupancy, I examined the expression of enhancer RNAs in cells shortly after removal of Oct4. Again, despite the increase in Nanog binding at the regulatory elements of pluripotency genes, and in line with the reduced level of H3K27ac at some loci, there was also a rapid decrease in the level of some enhancer RNA transcripts (Figure 4.15).

Figure 4.15

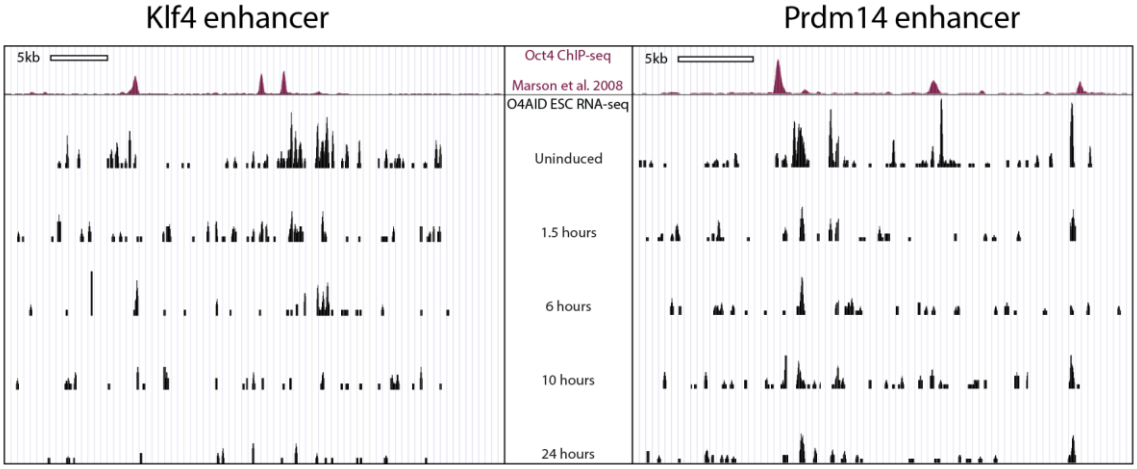



Figure 4.15 Changes in enhancer RNA expression following Oct4 depletion in Oct4-AID ESCs
Visualization of enhancer RNA expression at the Klf4 and Prdm14 enhancers in Oct4-AID ESCs following Oct4 depletion by RNA-seq. Oct4 ChIP-seq data from Marson et al., 2008 (publically available at GEO, accession GSE11724), are shown to indicate enhancer elements. The enhancer elements are centred ~52kb and ~35kb downstream of from the 3′ end of the Klf4 and Prdm14 loci respectively. For each locus, the y-axis represents library size normalised read density. Oct4-AID ESCs were maintained and induced in naïve-specific N2B27-2iL to minimize perturbations and were permanently maintained in the presence of dox.

## 4.3. Discussion

The phenomenon of pluripotency has been of great interest for many years, perhaps since it was first shown that cells taken from a donor blastocyst could integrate into a recipient blastocyst and contribute to normal embryonic development (Gardner, 1968). Since then, we have learned how to capture pluripotency *in vitro* (Evans and Kaufman, 1981; Martin, 1981), and how to maintain cells in a state equivalent to the preimplantation epiblast (Boroviak et al., 2015, 2014; Ying et al., 2008). Oct4 was perhaps the first factor described as essential for the establishment of the naïve epiblast (Nichols et al., 1998). Since then it has been shown that Oct4 is dispensable for initial generation of the

blastocyst structure, but required for segregation of the inner cell mass into naïve epiblast and primitive endoderm (Le Bin et al., 2014). We have also learned that Oct4 is essential for maintaining naïve pluripotency *in vitro* (Niwa et al., 2000). Despite this, the manner in which Oct4 is required for these processes is unclear; no downstream target, exogenous expression or knockout of which can rescue the Oct4 knockout phenotype, has been identified (Hall et al., 2009; Matoba et al., 2006; Niwa et al., 2000).

### 4.3.1. Conventional systems for depletion of Oct4 pass through an Oct4-low state

In order to identify essential targets of Oct4, I wanted to examine initial transcriptional changes in response to Oct4 depletion. It is notable that similar experiments have been performed before (Hall et al., 2009; Matoba et al., 2006) and have failed to yield an essential role for Oct4, but I anticipated that studying a larger number of timepoints would reveal more specific targets that could be further characterized. This system did reveal some interesting results, such as the rapid upregulation of several key trophectodermal transcription factors prior to downregulation of many pluripotency genes (Figure 4.2). It is tempting to interpret this as evidence that Oct4 is primarily required for the repression of trophectodermal genes, and following their depression these lead to loss of pluripotency and adoption of an extraembryonic fate. However, I wanted to rule out the possibility that my results were skewed as a result of cells passing through an Oct4-low state. It is known that cells exhibiting a low level of Oct4 can sustain expression of pluripotency markers under mild differentiation conditions (Karwacki-Neisius et al., 2013; Radzisheuskaya et al., 2013), and I wondered if transcription of such factors was being artificially maintained by the low levels of Oct4 that remain until all Oct4 protein eventually gets degraded. The half-life of Oct4 is sometimes reported to be very long (Lin et al., 2012; Wei et al., 2007); it took about a day for Oct4 to be fully removed from the cells used in this work (Figure 4.1c), and published work shows more than 10 hours for complete Oct4 depletion using a TetOFF system (Hall et al., 2009; Niwa et al., 2000). This is a significant amount of time, especially when one is trying to study direct and immediate responses to loss of Oct4. No timescale has been reported for cells to start showing responses to the Oct4-low state, and it is unclear what represents direct changes caused by lower Oct4 levels and what are compensatory changes induced by feedback loops or downstream responses to altered transcriptional networks.

Utilizing advances in imaging and counting mRNA molecules in single cells, it has recently been described that transcription occurs in bursts in prokaryotes (Cai et al., 2006; Golding et al., 2005) and

eukaryotes (Chubb et al., 2006), including in mammals (Raj et al., 2006). These bursts are characterized by pulses of polymerase activity interspersed with inactive periods. The length of these intervals between transcription are highly variable between genes, but are typically in excess of an hour and up to several hours (Suter et al., 2011). Reports suggest that the 'off' time for Nanog is between 1 and 4.5 hours (Muñoz Descalzo et al., 2013; Skinner et al., 2016), which would allow ample time for low Oct4 levels to affect Nanog burst frequency or size using conventional Oct4 depletion systems.

In order to avoid this confounding effect, I chose to establish new cell lines which utilise rapid depletion of Oct4 at the protein level (Figure 4.4a). It has previously been reported that the auxin inducible degron can reduce the half-life of tagged proteins to ~20 minutes in mammalian systems (Holland et al., 2012; Nishimura et al., 2009), and indeed I found that tagged Oct4 protein was greatly reduced within half an hour and almost undetectable within an hour of addition of IAA (Figure 4.5d and Figure 4.6d).

### 4.3.2. DNA binding of pluripotency factors is affected immediately following loss of Oct4

After ensuring that both the AID domain and Oct4 were unaltered in their function in this fusion protein (Figure 4.5 and Figure 4.6), I proceeded to re-examine the immediate effects of loss of Oct4. Since depletion of Oct4 was so rapid, I was able to examine the effect of loss of Oct4 prior to changes in the protein level of other pluripotency transcription factors. The interaction between Oct4 and Sox2 has been relatively well described; a model of the Oct4-Sox2-DNA complex has been generated via homology to an Oct1-Sox2-DNA crystal structure (Reményi et al., 2003; Yesudhas et al., 2016), and the cooperative binding of these proteins on DNA has been studied extensively *in vitro* (Ambrosetti et al., 2000, 1997; Reményi et al., 2003; Tokuzawa et al., 2003). These previous findings suggest that in the absence of Oct4, Sox2 should not interact as strongly at their co-binding sites. That said, single-molecule studies into the kinetics of Oct4 and Sox2 binding to DNA have suggested that Oct4 only slightly increases the residence time of Sox2 at specific binding sites (Chen et al., 2014). Despite this, I observed a decrease in Sox2 binding at the Sox2 regulatory element following loss of Oct4 (Figure 4.9b). As Biddle et al. suggest, there may be significant diversity in the cooperativity of Oct4 and Sox2 across the genome (Biddle et al., 2019). Further studies into genome-wide changes in Sox2 binding following acute loss of Oct4 could be helpful in further understanding the interactions of these two factors. It should be noted that these results are not very robust, due to sub-optimal ChIP conditions and lack of biological replicates. Nonetheless, to the best of my knowledge, this represents the first

time it has been shown that Oct4 is required for efficient binding of Sox2 to DNA *ex vivo* in mouse ESCs.

ChIP-qPCR and ChIP-seq against Nanog protein yielded an interesting and surprising result; in the absence of Oct4, greater levels of Nanog were found bound to the genome (Figure 4.9c, Figure 4.10b, Figure 4.11 and Figure 4.12). Remarkably, this extended beyond Oct4 co-bound sites, with a significant increase in Nanog levels observed at many binding sites lacking Oct4 (Figure 4.12a and b). While it seems plausible that Oct4 may physically occlude Nanog binding sites where they colocalise, resulting in competitive binding, this is somewhat contradicted by previous reports of simultaneous binding of Oct4 and Nanog to regulatory regions as detected by sequential ChIP (Medeiros et al., 2009). It is noteworthy, however, that while there is significant overlap between Oct4 and Nanog binding sites, the two proteins are typically separated by several nucleotides and may not physically interact on the DNA (Ferraris et al., 2011). To the best of my knowledge, it has not been shown whether Oct4 and Nanog interact cooperatively or competitively on DNA through *in vitro* assays, rather assumed from indirect evidence such as overlap in binding sites as described above.

On the other hand, it has been demonstrated that Nanog and Sox2 strongly interact, even in the absence of DNA (Gagliardi et al., 2013). Despite the fact that Oct4 and Sox2 are commonly presented as important co-binding factors, Sox2 actually co-localises more frequently with Nanog than Oct4 in genome-wide studies (Chen et al., 2008; Göke et al., 2011). It would be very interesting to look globally at changes in Sox2 binding in the immediate aftermath of Oct4 withdrawal; it may be that Sox2 is liberated from Oct4 co-binding sites, and recruits other transcription factors including Nanog to non-Oct4 bound sites. A previous study examined the localisation of Nanog and Sox2 24 hours after transcriptional depletion of Oct4 (King and Klose, 2017). In this work, they found that Sox2 and Nanog binding was reduced at many Oct4 co-bound sites correlating with reduced chromatin accessibility in the absence of Oct4. However, a subset of co-bound sites retained chromatin accessibility after Oct4 depletion, and at these sites binding of both Nanog and Sox2 was increased. Of course, this carries the caveat that such changes may be due to subsequent transcriptional and epigenetic changes downstream of loss of Oct4 rather than direct effects, but this could represent a similar redistribution of Sox2 and Nanog to sites where Oct4 binding is of less functional relevance.

It is currently unclear how the global chromatin state changes following acute depletion of Oct4. Previous work has suggested that Oct4 may act as a pioneer factor, opening chromatin that would

otherwise be highly compacted (King and Klose, 2017; Soufi et al., 2015). As a result, removal of Oct4 might result in the eviction of other transcription factors and repression of gene expression. Additionally, there could be significant changes in the chromatin architecture due to altered topological domains in the absence of Oct4. Decommissioning of super enhancers, many of which are bound by Oct4 in ESCs (Whyte et al., 2013), could lead to a reorganisation of the nuclear structure as these elements can interact with multiple promoter regions, even over large distances (Novo et al., 2018). It would be interesting to examine the chromatin state of cells before and after depletion of Oct4 to test these hypotheses. DNA FISH approaches could be used to examine the interactions of individual loci, while techniques such as Hi-C and Promoter-Capture Hi-C could be used to build a global picture of chromatin interactions and particularly promoter-superenhancer interactions (Novo et al., 2018). Other approaches such as MNase-seq and ATAC-seq could provide information about chromatin accessibility to determine whether chromatin is more highly compacted overall, and specifically in the region of Oct4 binding sites.

In agreement with existing literature, it appears that the half-life of Nanog is increased in the presence of Oct4 (Muñoz Descalzo et al., 2013). As a result, increased Nanog stability is not likely to be responsible for the increased genomic occupancy I observed. However, it would be interesting to see what other attributes of Nanog protein are altered in the absence of Oct4. It has been suggested that phosphorylation of Nanog by Erk1 results in ubiquitylation and subsequent degradation (Kim et al., 2014), a process antagonised by the de-ubiquitylase USP21 (Jin et al., 2016; Kwon et al., 2017; Liu et al., 2016). Nanog phosphorylation at partially overlapping sites has also been associated with an increase in stability via interaction with the prolyl isomerase Pin1 (Moretto-Zita et al., 2010). It would therefore be interesting to examine whether post-translational modification of Nanog is affected by the presence/absence of Oct4, to see whether Oct4 plays a role in recruiting these factors or protecting Nanog from their binding. It would also be interesting to see if the residence time of Nanog on DNA is altered in the absence of Oct4. It is becoming increasingly common to use single molecule tracking techniques to calculate the kinetics of the interactions between transcription factors and chromatin (Hemmerich et al., 2011; Loffreda et al., 2017; Mazza et al., 2012; Morisaki et al., 2014). This could indicate whether Nanog interacts more strongly with DNA in the absence of Oct4, possible explaining the increase in occupancy we observe. It could also reveal whether the increased binding we observe

via ChIP-seq is caused by an increase across all cells, or a reduction in the heterogeneity of Nanog binding between cells (Mueller et al., 2013).

### 4.3.3. Oct4 is required for transcription of key pluripotency factors

As previously established, conventional systems used to study loss of Oct4 are characterized by transit through an Oct4-low state. It has previously been shown that this can result in altered responses to signalling molecules and differential expression of several key pluripotency associated genes (Karwacki-Neisius et al., 2013; Radzisheuskaya et al., 2013). Therefore, while thousands of differentially expressed genes have been identified following depletion of Oct4 (Hall et al., 2009; Matoba et al., 2006), it is unclear what are genuine responses to removal of Oct4 rather than responses to reduced levels of Oct4. Consequently, the rapid Oct4 depletion system is well suited to examining the immediate impacts in a way that existing systems cannot.

As I have described, expression of several key pluripotency factors drops rapidly as a result of loss of Oct4 (Figure 4.15a-c). Some of these factors show contrasting responses to acute versus prolonged depletion of Oct4; for instance, in my LoxCre system Klf4 responds slowly to removal of Oct4, whereas upon rapid loss of Oct4 it is quickly downregulated. Similarly, by microarray analysis Nanog was not differentially expressed on transcriptional inhibition of Oct4 (Hall et al., 2009), while transcript levels dropped prior to substantial trophoblast gene expression in my system (Figure 4.15a and ).

Given that Nanog binding is sustained at these loci, these data suggest that Oct4 binding may be required for the expression of a subset of pluripotency genes. It would be beneficial to explore this more directly by examining changes in nascent transcript levels in the immediate aftermath of acute Oct4 depletion. Interestingly, expression of enhancer RNAs at some pluripotency loci is decreased rapidly on loss of Oct4 (Figure 4.7). Transcription of enhancer RNAs is strongly correlated with enhancer activity (Azofeifa et al., 2018; De Santa et al., 2010; Kim et al., 2010), so this finding implies that some Oct4-bound enhancers rapidly lose their functionality when Oct4 is removed.

These differences go some way to explaining past discrepancies in the perceived role of Oct4 in the pluripotency network. As Hall *et al.* note, Nanog has been described as an Oct4 target gene, and there is significant biochemical data validating that Oct4 binds to the Nanog enhancer region and is required for its expression (Kuroda et al., 2005; Rodda et al., 2005), yet Oct4 depletion appeared to have a limited, even positive impact on Nanog mRNA. I propose that the reason for this is that while Oct4

may be essential for Nanog expression, above a minimal level Oct4 negatively regulates the binding of other transcriptional activators. Thus, as Oct4 levels drop and cells pass through an Oct4-low state, binding of Oct4 to the Nanog enhancer is reduced and other transcription factors can be recruited, buffering Nanog expression levels. Once Oct4 is entirely depleted, very rapidly in the case of induced degradation, Nanog expression falls as the locus is no longer transcribed. Clearly this is highly speculative, and requires a significant amount of work to validate. However, this would reproduce the phenotypes observed with varying Oct4 levels (Figure 4.16): a minimal level would be required for expression of key pluripotency factors, and below this cells would not be able to maintain pluripotency; above this, a low level of Oct4 would be optimal for self-renewal, with strong expression of many pluripotency factors; higher, at an ESC level, Oct4 would begin to show a repressive effect on genes such as Nanog, generating a state which is permissive for both self-renewal and differentiation; finally, on overexpression of Oct4, transcription factors would be excluded from the regulatory regions of key pluripotency genes, resulting in collapse of the naïve network and induction of differentiation.

Figure 4.16

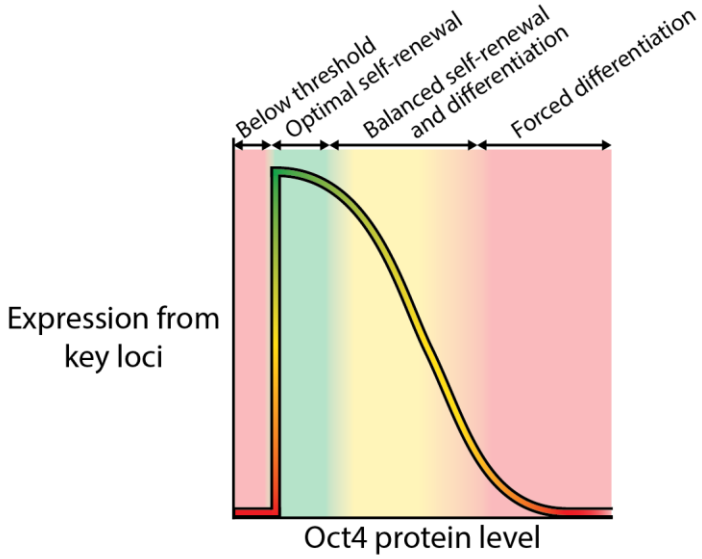



Figure 4.16 A model of the relationship between Oct4 protein level and expression of a set of pluripotency genes
A graph showing the relationship between the level of Oct4 protein in a cell and expression from key pluripotency loci with the corresponding effect on cellular identity. At very low levels, Oct4 is below the threshold required for efficient expression of pluripotency genes. A slight increase in Oct4 protein level results in strong expression of these key genes, resulting in optimal self-renewal. At this level, differentiation is inhibited as a result of the strong expression of naïve-specific genes. Cells expressing a higher level of Oct4, similar to the level seen in ESCs, have weaker expression of key pluripotency genes due to effects such as displacement of Nanog protein from regulatory elements. In such cells, there is a balance between self-renewal and differentiation; depending on culture conditions, cells can be maintained or induced to exit pluripotency. At even higher levels, Oct4 reduces expression of key pluripotency factors even further, to such a point that the naïve network is no longer stable and self-sustaining. As a result, these cells can no longer retain their pluripotent identity and are forced to differentiate.

Under this model, there would be several key pluripotency factors which require Oct4 for their expression. Thus, co-expression of several exogenous genes would be required to rescue the Oct4 knockout phenotype. Unfortunately, the data generated so far are not statistically rigorous, preventing the generation of a full list of candidates, though likely factors include Nanog, Esrrb, Klf4 and Lifr, which are rapidly downregulated upon proteolytic Oct4 loss, and in some cases on conventional Oct4 depletion.

### 4.3.4. Exit from pluripotency following loss of Oct4 follows a non-conventional trajectory

Typically, exit from pluripotency occurs in stages as cells pass through progressively more differentiated states, until eventually they undergo fate restriction and lose their potential to form other lineages. *In vivo*, cells of the naïve pre-implantation epiblast undergo polarization as the blastocyst implants into the host (Bedzhov and Zernicka-Goetz, 2014). At the same time, these cells undergo global chromatin remodelling switching to a much less permissive state (Auclair et al., 2014; Smith et al., 2012; Surani et al., 2007), with coincident transcriptional rewiring and exit from the naïve state (Boroviak et al., 2015, 2014). The cells of the post-implantation epiblast rely on different external signals and internal transcription factors and enhancer usage to promote survival and proliferation. As the surrounding extraembryonic tissues mature they polarize the epiblast, which begins to specify the fate of cells in a regional manner (Tam and Loebel, 2007). Subsequently, as gastrulation occurs and presumptive mesendoderm cells pass through the primitive streak, they differentiate into one of the three germ layers (Tam and Loebel, 2007). Within about one day of this, around the time of somitogenesis, all cells have lost their capacity for pluripotency (Damjanov et al., 1971; Osorno et al., 2012).

These events are recapitulated *in vitro* through the differentiation of ESCs, where they have been examined in molecular detail. ESCs maintained in 2iL culture conditions are equivalent to the E4.5 naïve epiblast (Boroviak et al., 2015; Silva et al., 2008; Ying et al., 2008). As these cells are subjected to differentiation-inducing stimuli, they exit naïve pluripotency via downregulation of naïve-specific transcription factors and upregulation of markers of the post-implantation epiblast. Initially, cells appear to pass through a pass through a state more advanced than naïve pluripotency yet more permissive than primed, with competency for germ cell differentiation (Hayashi et al., 2011; Kalkan et al., 2017; Kalkan and Smith, 2014; Mulas et al., 2017). Subsequently, cells progress to the primed state of pluripotency (Brons et al., 2007; Buecker et al., 2014; Joo et al., 2014; Tesar et al., 2007; Yang et al.,

2014), a heterogeneous state with fluctuating expression of lineage markers and similarities to the primitive streak *in vivo* (Kojima et al., 2014; C. Liu et al., 2018; Song et al., 2016; Tsakiridis et al., 2014). Further differentiation involves downregulation of general and primed pluripotency factors and upregulation of lineage marks as cells become committed (Semrau et al., 2017; Trott and Martinez Arias, 2013).

Coincident with the collapse of the naïve network in Oct4-AID ESCs, expression of general pluripotency factors begins to drop. There is no upregulation of primed pluripotency genes in these cells, nor are lineage markers expressed. Instead, cells directly start to upregulate silenced extraembryonic genes (). However, this makes sense in the context of previous observations that Oct4 plays an important role in exit from the naïve state (Karwacki-Neisius et al., 2013; Radzisheuskaya et al., 2013) and is recruited to different regulatory elements to drive differentiation (Yang et al., 2014).

Notably, published microarray data from TetOFF Oct4 cells show downregulation of naïve pluripotency genes accompanied by extended expression of general pluripotency genes such as Sall4 and Nanog and upregulation of some post-implantation genes like Otx2 and Sox4, but with a failure to upregulate others including Fgf5 and Lef1 (Hall et al., 2009). It is therefore striking that exit from pluripotency follows such a radically different trajectory following acute depletion of Oct4.

### 4.3.5. Exploring the mechanistic properties of Oct4

Together, this body of work demonstrates that there is an absolute requirement for a minimal level of Oct4 for expression of a subset of factors that help to maintain the naïve state (Figure 4.7). This is despite an increase in Nanog binding to the regulatory elements of many such genes (Figure 4.9-12). The fact that a threshold level of Oct4 is required for the expression of a set of core pluripotency genes, and the observation that binding of the strong pluripotency agonist Nanog to the genome is globally lower in cells expressing Oct4 than in cells depleted for Oct4 suggests a coherent mechanistic explanation for the various phenotypes associated with different levels of Oct4. In Oct4-low cells, Oct4 binds and activates transcription of pluripotency-associated genes, resulting in the establishment of a robust network that is relatively insensitive to differentiation cues. As levels of Oct4 increase, factors such as Nanog are excluded from the genome, resulting in weaker activation of core pluripotency genes, until at sufficiently high Oct4 levels pluripotency can no longer be maintained. Meanwhile, in the absence of Oct4 these genes cannot be activated and so the naïve state can neither be induced nor maintained (Figure 4.16).

# Chapter 5: Exit from naïve pluripotency induces expression of Xist, which is misregulated in extraembryonic differentiation

## 5.1. Introduction

X-chromosome inactivation is a key developmentally-regulated process in mice. It is well established that female cells exiting naïve pluripotency inactivate one randomly-selected X-chromosome in order to balance X-linked gene expression between sexes. The long non-coding RNA Xist is transiently expressed from both X-chromosomes prior to being silenced and one chromosome and highly upregulated on the other. It coats the chromosome from which it is expressed *in cis*, recruiting factors which silence and inactivate this chromosome. It has previously been established that depletion of pluripotency factors induces expression of Xist (Navarro et al., 2008), demonstrating a role of for these components in the regulation of this key inactivation factor; however, ablation of the primary pluripotency-factor-bound Xist regulatory element does not interfere with either XCI or reactivation of the silenced X chromosome in the establishment of naïve pluripotency (Minkovsky et al., 2013).

Under the conventional model of initiation of XCI, in early differentiation cells undergo a process of identifying their X chromosomes through "sensing", counting their number relative to autosomes, then cells with multiple X chromosomes (i.e. females for euploid cells) choose which of these are to undergo inactivation, and initiate inactivation through upregulation of Xist which subsequently spreads along the X chromosome, coating it and inducing transcriptional silencing (Boumil and Lee, 2001; Goto and Monk, 1998; Starmer and Magnuson, 2009) (Figure 5.1a).

### 5.1.1. A new model for XCI

Figure 5.1

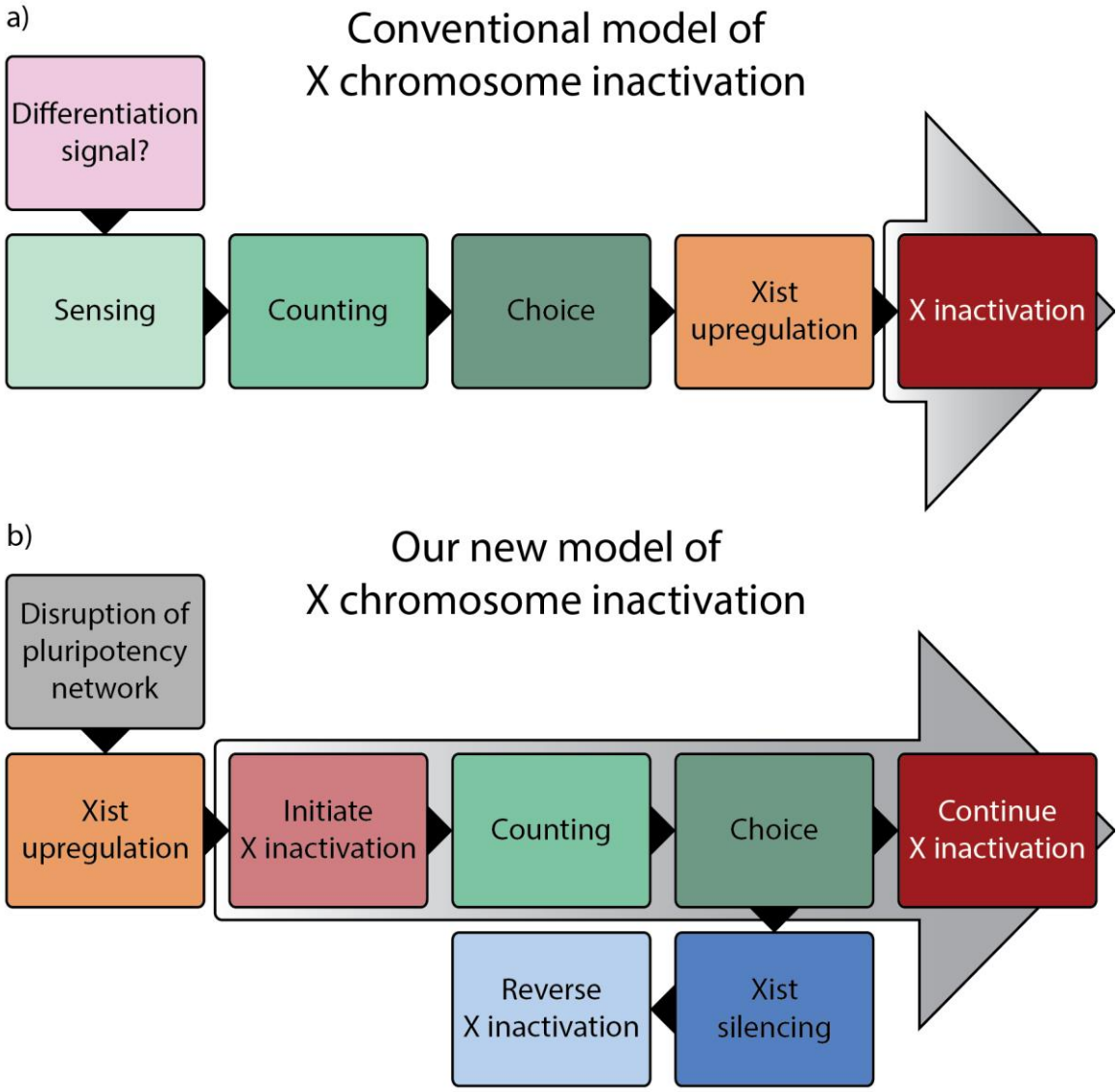



Figure 5.1 Schematic of the conventional model for XCI and our prospective model for XCI
a) The conventional model of XCI. An unknown signal in differentiating ESCs causes sensing of X chromosomes and counting. If XCI is required, the choice of which chromosome(s) to inactivate is made, followed by upregulation of Xist which induces XCI. b) Our model of XCI. The collapse of the naïve transcription network results in upregulation of Xist in all differentiating cells, resulting in the initial stages of XCI. Subsequently, cells undergo counting and choice. Xist is silenced on chromosomes that are not destined for XCI, allowing the early inactivation events to be reversed. Meanwhile, XCI is completed on chosen chromosomes.

Data from our group have suggested that initially Xist expression is induced by the collapse of the naïve network, regardless of gender. In studying this effect, Elsa Sousa observed that the upregulation of Xist in male and female cells appears to occur after downregulation of pluripotency markers and prior to upregulation of differentiation genes, suggesting that a reduction in the level of naïve factors could be responsible for de-repression of Xist, leading to its initial upregulation; counting and choice

may be later events that lead to silencing of Xist on future active chromosomes rather than activation of Xist from presumptive inactive chromosome (Sousa et al., 2018) (Figure 5.1b). This is based on observations *in vivo* and *in vitro* that Xist is transiently upregulated on exit from pluripotency in males and is initially expressed biallelically in females, though it is unclear whether this occurs in all cells or a subpopulation. Expression in males initially matches that in females; however, it only reaches a small fraction of the level attained in females as they continue to accumulate Xist RNA, whereas transcription is rapidly shut down in males (Sousa et al., 2018).

Notably, upon deletion of Nanog from male or female mouse ESCs, Sousa et al. observed upregulation of Xist over a time period in which several other pluripotency markers maintained their expression. This suggests a very direct role for some pluripotency factors in the repression of Xist in naïve cells. In order to expand our understanding of this phenomenon, I chose to examine the effect of loss of Oct4 on expression of Xist.

## 5.2. Results

In order to test whether repression of Xist is a general property of pluripotency factors, I decided to examine the response of Xist on loss of Oct4. Unfortunately, the presence of the antisense Tsix lncRNA prevents analysis of Xist expression from non-strand specific RNA-seq data; I therefore repeated the Oct4 deletion timecourse using a conventional Lox-Cre system and analysed Xist expression using RT-qPCR with a probe that spans the Xist exon 1-2 boundary to achieve specificity. This also had the advantage of allowing me to examine the response in both male and female cell lines derived from littermates; ESCs were previously derived from male and female $Oct4^{F/-}$ blastocysts that constitutively express CreER from a Rosa26 knock-in allele. Surprisingly, regardless of whether male or female ESCs were used, Xist was upregulated to levels seen in female somatic cells, rather than the limited level of upregulation observed in normal differentiating male ESCs or on loss of Nanog (Figure 5.2) (Sousa et al., 2018); in normal differentiation, expression of Xist is only transiently upregulated to very low levels in male cells (Sousa et al., 2018). Even more interestingly, expression of Xist was sustained for at least 7 days on removal of Oct4, whereas Xist is only expressed for ~24-48 hours on conventional differentiation of male ESCs. In male trophoblasts Xist is entirely silenced by imprints (Goto and Takagi, 2000; Latham, 1996; Lee, 2000), so it is somewhat surprising that it is so strongly upregulated on deletion of Oct4. Examination of two long non-coding RNA regulators of Xist, Tsx and Tsix,

showed that these were both rapidly downregulated on removal of Oct4 (Figure 5.2), unlike regular differentiation in which both of these factors are transiently upregulated (Sousa et al., 2018)

Figure 5.2

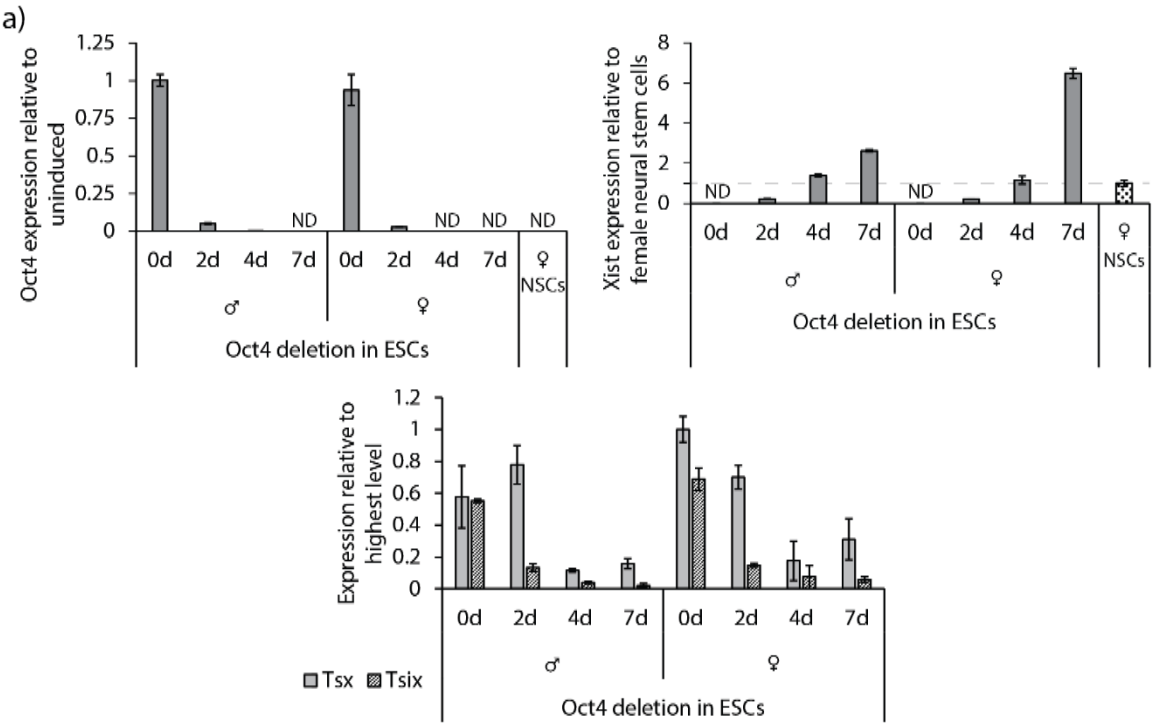



Figure 5.2 Xist transcription is induced to high and sustained levels in male ESCs following depletion of Oct4
RT-qPCR analysis of Oct4 (left), Xist (right), and Tsx and Tsix (bottom) expression in male and female Oct4F/- CreERt2 ESCs following genetic ablation of Oct4 by addition of 4-OHT and female neural stem cells (NSCs). Female NSCs are used to indicate the level of Xist expression in female somatic cells. Values represent the mean of three replicate qPCR reactions and error bars represent ± SD; values were calculated by the ΔΔCt method using the housekeeping gene Gapdh as a reference gene; Oct4 expression was normalized to uninduced male Oct4$^{F/-}$ Rosa26$^{+/CreER}$ ESCs, Xist expression was normalized to female NSCs, Tsx and Tsix expression were normalized to highest value. One representative male and female Oct4$^{F/-}$ Rosa26$^{+/CreER}$ ESC line are shown of two male and two female littermate-derived ESC lines tested. ESCs were maintained in naïve-specific N2B27-2iL conditions then switched to differentiation-permissive SL conditions on addition of 4-OHT.

Trophoblast stem cells (TSCs) can be maintained *in vitro* by plating trophectoderm tissue in chemically defined media. In order to test whether cells could self-renew with such high levels of Xist, I differentiated male Oct4$^{F/-}$ ESCs by transferring them to TSC maintenance conditions at the same time as inducing deletion of Oct4 with 4-OHT. I was able to generate a TSC line (Figure 5.3a); however, every passage was accompanied with extreme cell death, and cells had to be transferred into a fresh well without removing any material in order to maintain the line. These cells exhibit a TSC-like morphology (Figure 5.3a) and express the marker Elf5, and to a lower level Cdx2 (Figure 5.3b). They have fully downregulated the ESC marker Rex1. However, they proliferated slowly, only requiring passaging every 8-9 days. Xist appears to have been strongly downregulated in these cells (Figure 5.3b), perhaps suggesting that cells that continued to express Xist were outcompeted or died.

Figure 5.3

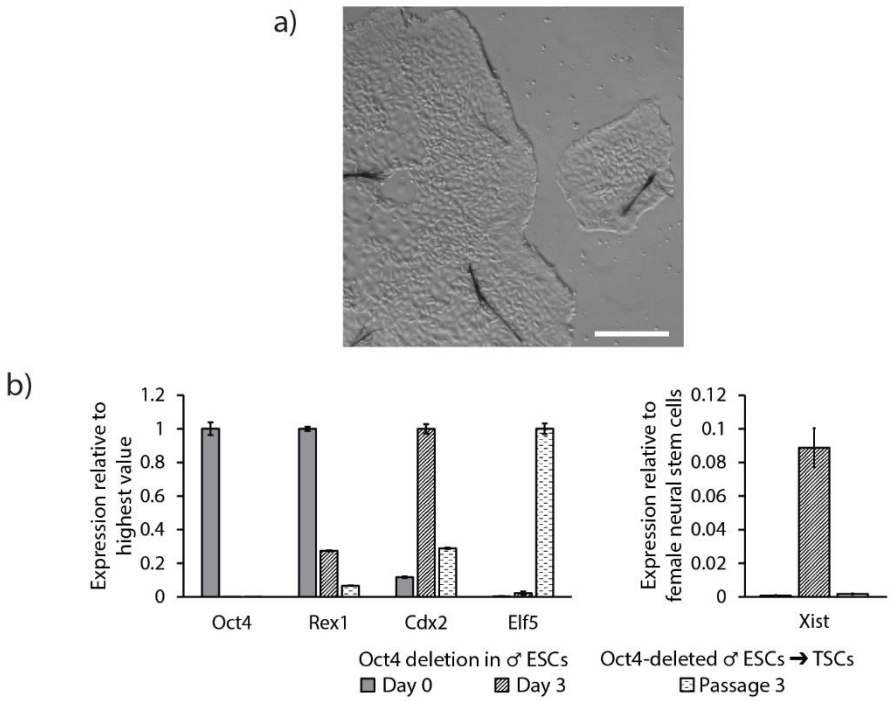



Figure 5.3 Following deletion of Oct4 in ESCs, trophoblast stem cells can be captured *in vitro*
Male ESCs can be captured *in vitro* following removal of Oct4. Male Oct4$^{F/-}$ Rosa26:CreER were maintained in N2B27-2iL; Oct4 deletion was induced by addition of 4-OHT, and at the same time the growth medium was changed to TSC media. a) Phase contrast image of cells following deletion of Oct4 and transition to TSC media after one passage. Crystals are the result of precipitation of XAV939 from the media. Scale bar represents 250µm. b) RT-qPCR analysis of gene expression in male Oct4$^{F/-}$ Rosa26:CreER cells following addition of 4-OHT. Cells were maintained and induced in N2B27-2iL (Oct4 deletion in male ESCs) or induced in TSC media (Oct4-deleted male ESCs → TSCs). Values represent the mean of three replicate qPCR reactions and error bars represent ± SD; values were calculated by the ΔΔCt method using the housekeeping gene Gapdh as a reference gene and normalized to the highest value, or to female neural stem cells in the case of Xist.

### 5.2.1. Generating cell lines to study XCI in extraembryonic differentiation from ESCs

In order to test whether this observation that Xist is strongly upregulated in male cells on depletion of Oct4 is a result of this unconventional exit from the naïve state or a peculiarity of this non-physiological disruption to the transcription factor network, it would be useful to examine the response of ESCs to alternative methods of induced extraembryonic differentiation. I therefore sought to establish systems to achieve extraembryonic differentiation in this manner, and examine the trajectory of differentiation and the impact that this has on the regulation of Xist.

It has previously been demonstrated that overexpression of certain transcription factors is sufficient to cause differentiation of mouse ESCs to extraembryonic fates (Blij et al., 2015; Fujikura et al., 2002; Kinoshita et al., 2015; Niwa et al., 2005; Shimosato et al., 2007). I designed dox-responsive constructs for inducible expression of Cdx2 and Gata6 in order to promote differentiation towards trophectoderm and extraembryonic endoderm respectively (Fujikura et al., 2002; Niwa et al., 2005). I

generated Piggybac plasmids harbouring these constructs and randomly integrated either Cdx2 or Gata6, as well as constitutively expressed rtTA3, into male mouse ESCs carrying a GFP reporter at the Rex1 locus (Figure 5.4). I validated the gender of the cell lines by profiling expression of X- and Y-linked gene expression; genes from both sex chromosomes where robustly expressed, demonstrating that the cells were indeed male (Figure 5.5a).

Figure 5.4

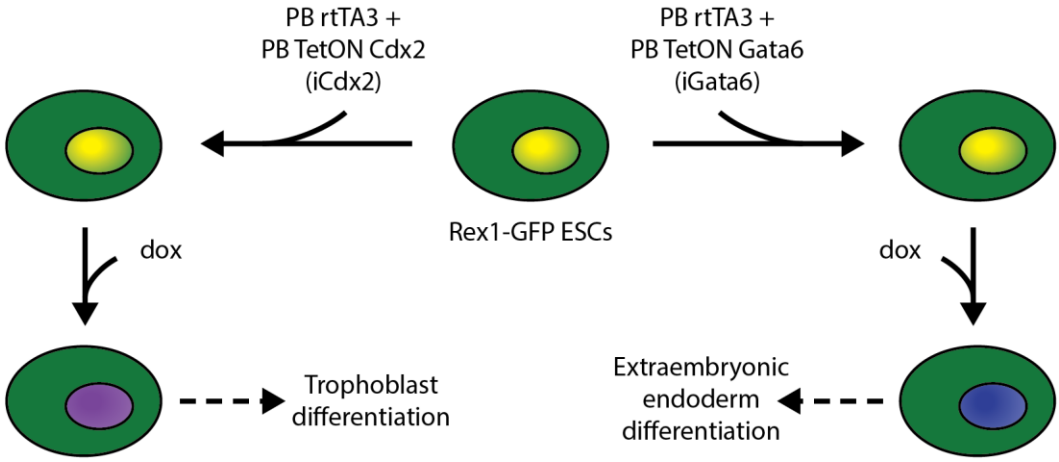



Figure 5.4 Experimental design to generate cell lines capable of induced extraembryonic differentiation
Male mouse ESCs were transfected with a randomly integrating, constitutively expressed rtTA3 (dox-inducible transactivator) construct and a randomly integrating TRE Cdx2 or TRE Gata6 cassette, together iCdx2 and iGata6 respectively. These cells were induced to differentiate to extraembryonic lineages by addition of dox towards trophoblast or extraembryonic endoderm identities respectively.

After selection, once the constructs had stably integrated, I tested that the transgenes induced appropriately and the cells differentiated to their appropriate lineages as expected. Inducible Cdx2 cells downregulated the naïve gene Nanog and upregulated the trophoblast marker Elf5 (Figure 5.5b, left), and flattened on addition of dox. After four days in culture, trophoblast giant cells also began to appear (not shown). Inducible Gata6 cells immediately adopted a characteristic cobblestone morphology, which became more pronounced as induction continued (not shown). They also lost Nanog expression and expressed the extraembryonic gene Gata4 (Figure 5.5b, right).

Figure 5.5

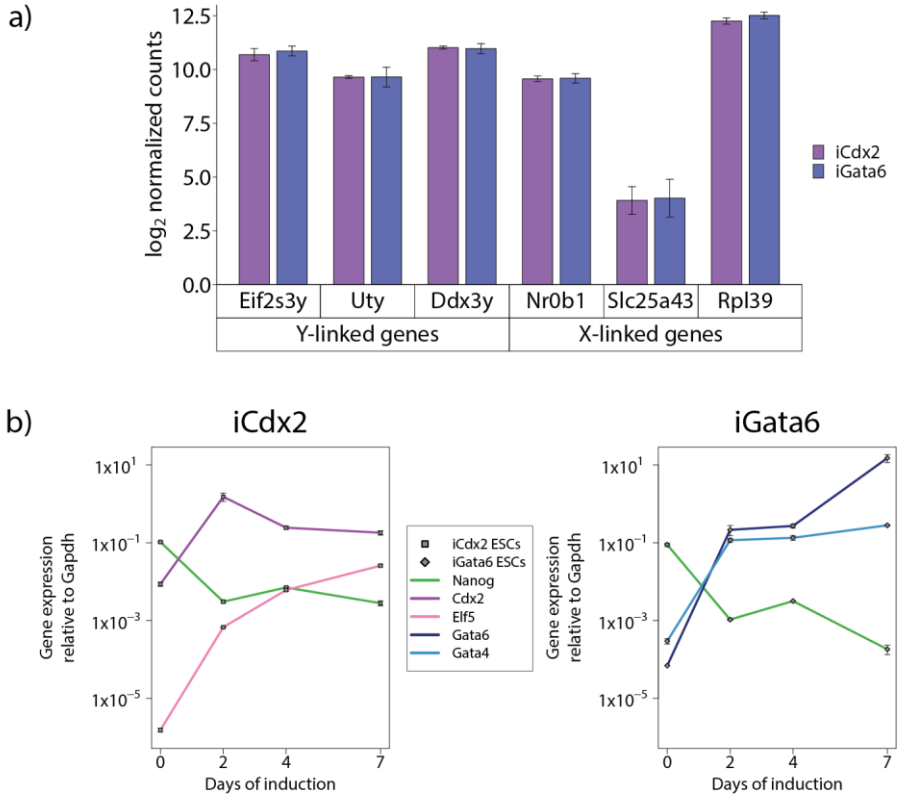



Figure 5.5 Characterizing cell lines for inducible extraembryonic differentiation
a) Male gender of iCdx2 and iGata6 ESCs was confirmed by measuring expression of X and Y-linked gene expression. Graph shows $\log_2$ library size normalized read counts of three X-linked and three Y-linked genes in uninduced cells maintained in N2B27-2iL, taken from RNA-seq data with normalization performed using DESeq2. Values represent the mean of three biological replicates, and error bars represent ± SD. b) RT-qPCR analysis of iCdx2 (left) and iGata6 (right) ESCs in a timecourse following dox induction of transgenes. Expression of the pluripotency factor Nanog, the induced transgene, and a representative extraembryonic differentiation marker are shown on a $\log_{10}$ scale. Values represent the mean of three replicate qPCR reactions and error bars represent ± SD; values were calculated by the ΔCt method using the housekeeping gene Gapdh as a reference gene. Cells were maintained in naïve-specific N2B27-2iL and transferred to differentiation-permissive SL on addition of dox.

### 5.2.2. Extraembryonic differentiation from male mouse ESCs results in Xist upregulation

Once I had confirmed that the cell lines responded as anticipated to dox induction, i.e. they exited pluripotency and differentiated towards extraembryonic lineages, I performed a timecourse and sent samples for strand-specific high throughput RNA sequencing. Analysis of these data shows that exit from pluripotency differs across different routes to extraembryonic lineages. iCdx2 cells continue to express naïve-specific and general pluripotency genes longer than iGata6 (Figure 5.6a and Figure 5.7a), although they are strongly downregulated. Both iGata6 and iCdx2 lines still show little activation of markers of primed pluripotency, however. Instead they directly activate the appropriate extraembryonic gene network following downregulation of naïve factors (Figure 5.6a, Figure 5.7a), with iCdx2 cells upregulating TSC markers such as Hand1, Elf5 and Cdh3, and iGata6 cells expressing

high levels of Gata4, Sox7 and Foxa2. Interestingly, induction of either Cdx2 or Gata6 results in a rapid diversion from the path of normal embryonic differentiation (Figure 5.6b).

Figure 5.6

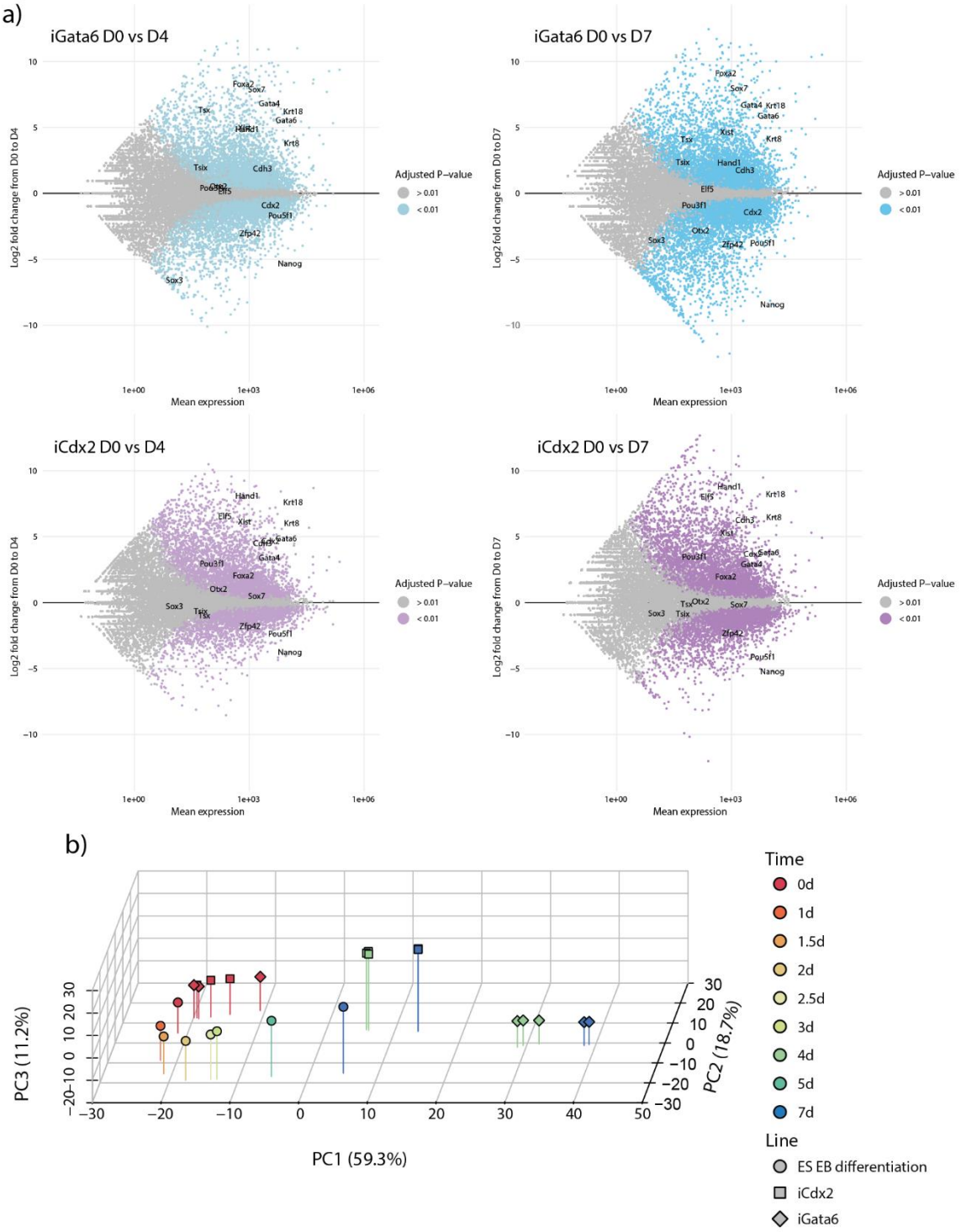


Figure 5.6 RNA-sequencing analysis of iGata6 and iCdx2 differentiation
a) MA plots showing each gene as a single point. The mean expression of each gene between day 0 and the indicated timepoint is shown on the x-axis, and the log2 fold change is shown on the y-axis; positive values indicate an increase in expression from day 0 to the indicated time point and negative values indicate a decrease. Genes showing a statistically significant change (adjusted P-value < 0.01) are coloured, and several key pluripotency, extraembryonic differentiation and XCI-associated genes are annotated. b) PCA plot showing the first three principal components of variance for the iCdx2 and iGata6 induction timecourses in addition to an embryoid body differentiation timecourse from ESCs (data from Elsa Sousa).

As with Oct4 knockout induced trophectoderm differentiation, overexpression of Cdx2 resulted in prolonged expression of Xist to a high level (Figure 5.6a, Figure 5.7a), validating this finding and demonstrating that it is not an artefact specifically caused by non-physiological Oct4 manipulation. Remarkably, Gata6-induced extraembryonic endoderm differentiation also caused significant transcription of Xist over the period of a week. This indicates that protracted Xist expression may be a general property of male cells differentiating to extraembryonic lineages.

I validated the induction of Xist following extraembryonic differentiation by RT-qPCR, using female neural stem cells to gauge the level of expression (Figure 5.7b). In the case of Cdx2 induction, Xist RNA was detected at approximately the same level as female somatic cells, while overexpression of Gata6 induced Xist around one fifth of this; this level of Xist expression without manipulation of the X inactivation centre in male differentiation is only paralleled in my previous Oct4 deletion experiments to the best of my knowledge.

Figure 5.7

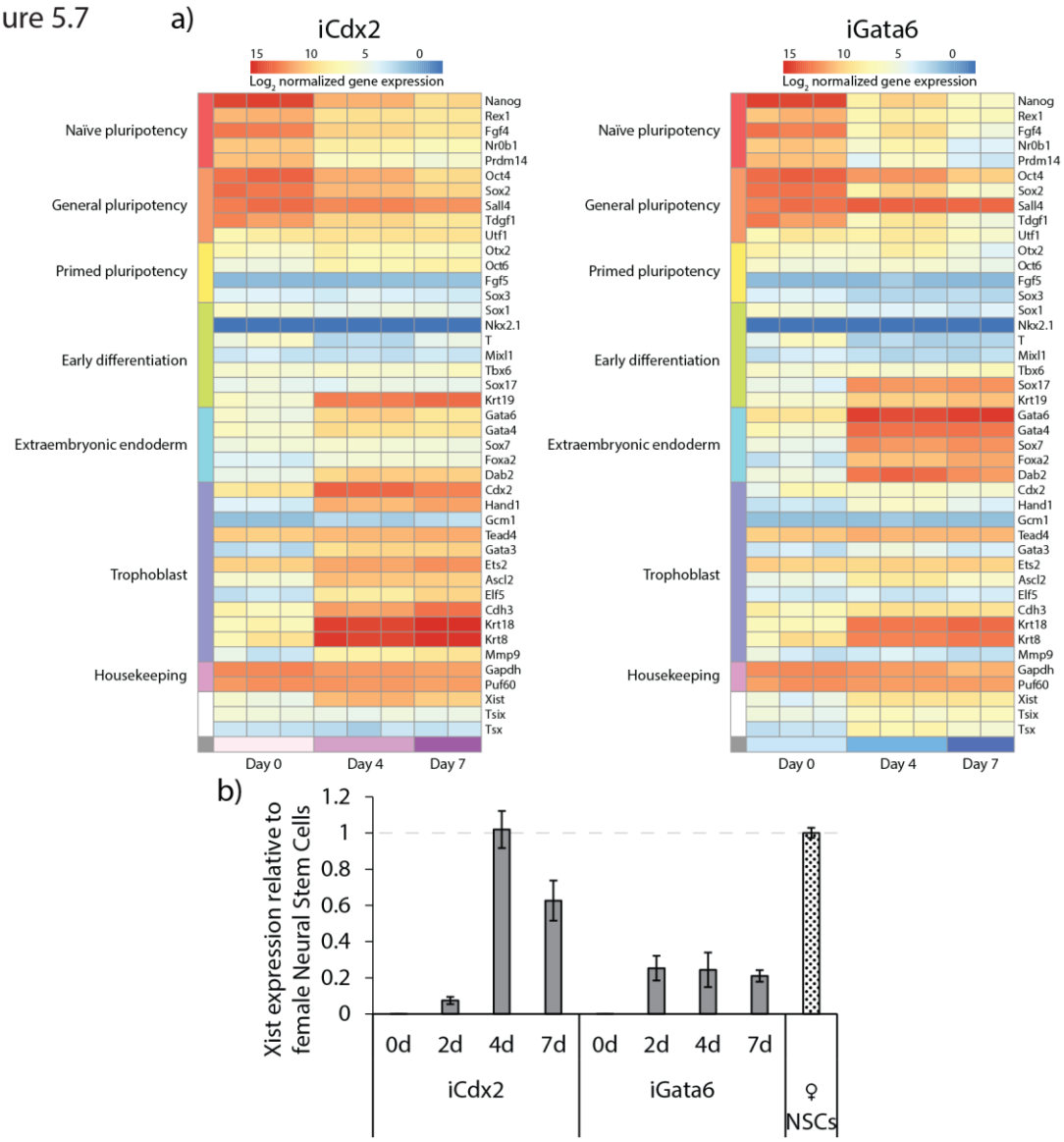



Figure 5.7 Expression analysis of iGata6 and iCdx2 cells
a) Heatmap showing expression of genes representative of several cell identities in iCdx2 (left) and iGata6 (right) ESCs over a timecourse following addition of dox. Three biological replicates are show for cells prior to induction and four days post-induction, and two biological replicates are shown for cells seven days post-induction. Intensities represent $log_2$ normalized expression calculated from RNA-seq data using DESeq2. b) RT-qPCR analysis of Xist in iCdx2 and iGata6 ESCs and female NSCs. Values represent the mean of three replicate qPCR reactions and error bars represent ± SD; values were calculated by the ΔΔCt method using the housekeeping gene Gapdh as a reference gene and normalised to female NSCs. ESCs were maintained in naïve-specific N2B27-2iL and transferred to differentiation-permissive SL on addition of dox.

It is possible to probe expression across the XIC using strand-specific RNA-sequencing. Elsa Sousa performed a timecourse of embryoid body differentiation from male ESCs and generated strand-specific RNA-seq data in order to demonstrate transient upregulation of Xist. In reviewing this data, we can see that embryoid body differentiation is associated with rapid upregulation of Tsix and Tsx, negative regulators of Xist (Figure 5.8). Interestingly, on induction of Gata6, Tsx and Tsix are only

gradually induced (Figure 5.8), which might restrain the upregulation of Xist without robustly extinguishing its expression as seen in embryoid differentiation. Meanwhile, neither Tsx nor Tsix are strongly upregulated on overexpression of Cdx2 (Figure 5.8). It appears that in the absence of sufficient induction of inhibitors, expression of Xist continues unimpeded.

Figure 5.8

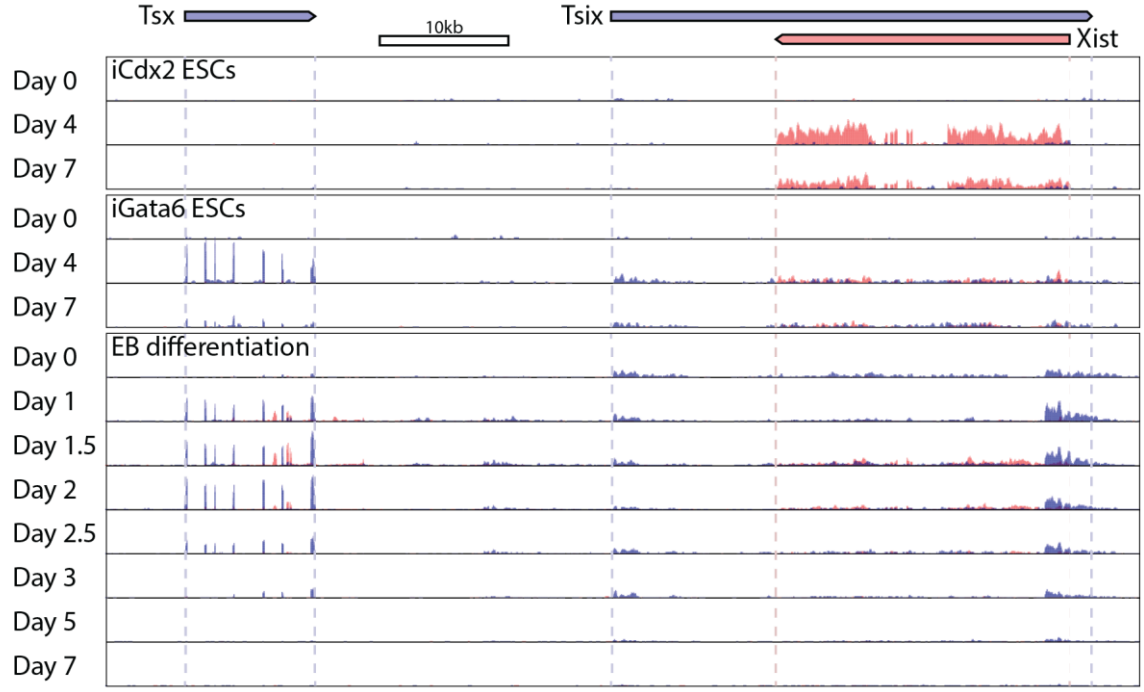



Figure 5.8 Expression within the XIC during extraembryonic and embryonic differentiation
Visualization of strand-specific RNA-seq data from male iCdx2 and iGata6 ESCs following dox induction and wild-type male cells during embryoid body (EB) differentiation. Strand-specific expression is indicated in blue (sense) and red (antisense) at Tsx, Tsix and Xist. iCdx2, iGata6 and wild-type ESCs were maintained in N2B27-2iL. iCdx2 and iGata6 ESCs were induced with dox in SL. Wild-type ESCs were plated in low-attachment dishes in serum-containing medium without LIF. The y-axis represents library size normalised read density, and scaled to the maximum expression for each cell line.
Embryoid body differentiation and corresponding RNA preparation was performed by Elsa Sousa, data are publically available at GEO, accession number GSE109173. Strand-specific alignment of RNA-seq data was performed by Sabine Dietmann and Michael Barber.

Oct4 binds in the region of Tsx and Tsix in ESCs, while other pluripotency factors such as Nanog only bind within intron 1 of Xist (Figure 5.9a). This is particularly relevant as Oct4 is far more rapidly downregulated following Cdx2 induction than following Gata6 induction or in normal embryonic differentiation (Figure 5.9b). It is plausible that the lack of Oct4 prevents induction of Tsx and Tsix, resulting in cells achieving high and sustained levels of Xist expression following Cdx2 overexpression or Oct4 depletion. Meanwhile, the sustained expression of Xist seen after Gata6 induction could relate to the absence of factors involved in the normal silencing of Xist in somatic differentiation, or changes in the Oct4 binding profile resulting in weaker induction of Tsx and Tsix.

Figure 5.9

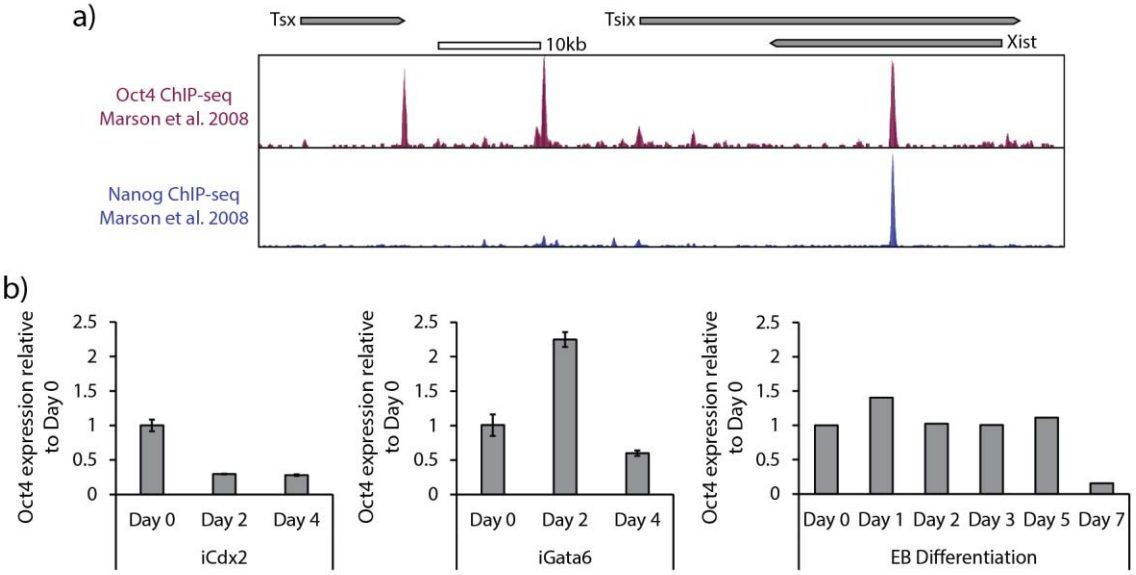



Figure 5.9 Oct4 expression in extraembryonic and EB differentiation, and Oct4 binding sites within the XIC
a) Visualization of Oct4 and Nanog binding within the XIC by ChIP-seq using data from Marson et al. 2008. b) Expression of Oct4 at early timepoints in iCdx2, iGata6 and EB differentiation. iCdx2 and iGata6 timecourses were analysed by RT-qPCR, Oct4 expression was quantified from RNA-seq data. For iCdx2 and iGata6, values represent the mean of three replicate qPCR reactions and error bars represent ± SD; values were calculated by the ΔΔCt method using the housekeeping gene Gapdh as a reference gene and normalised to day 0. For EB differentiation, values represent library size normalized read counts, normalized to day 0.
Embryoid body differentiation and corresponding RNA preparation was performed by Elsa Sousa, data are publically available at GEO, accession number GSE109173. Oct4 and Nanog ChIP-seq data are taken from Marson et al. 2008, and are publically available at GEO, accession number GSE11724.

### 5.2.3. Oct4 plays a direct role in repression of Xist

It is notable that in all of the experiments so far I have demonstrated strong and persistent upregulation of Xist in exit from pluripotency. ESCs differentiate when subjected to increased Oct4 levels (Niwa et al., 2000), unlike some pluripotency factors such as Nanog or Esrrb, which strengthen the naïve state when overexpressed (Chambers et al., 2003; Festuccia et al., 2012). Initial reports suggested that overexpression of Oct4 in ESCs primarily induces mesendodermal differentiation (Niwa et al., 2000; Thomson et al., 2011). However, it appears that this is highly dependent on the culture environment, with Oct4 overexpression leading to efficient neurectodermal (Shimozaki et al., 2003) and cardiac differentiation (Zeineddine et al., 2006) in appropriate conditions, and markers from all three germ layers are upregulated following induction of Oct4 in permissive conditions (Radzisheuskaya et al., 2013; Thorold W. Theunissen et al., 2011).

Given that I have demonstrated such strong induction of Xist on loss of Oct4, it seems plausible that overexpression of Oct4 could have an antagonistic effect. Additionally, male ESCs display persistent expression of Xist, albeit at a very low level, in differentiation permissive SL conditions. Oct4

overexpression is therefore an ideal system to test both whether Xist upregulation is a common feature of differentiation induced by transcription factor overexpression, and more excitingly, whether Oct4 is able to inhibit induction of Xist.

With this in mind, I took advantage of a previously established clonal TetON Oct4 cell line (iOct4 ESCs), generated by former colleague Aliaksandra Radzisheuskaya (Figure 5.10a). She transfected wild type male mouse ESCs in the same manner that I established iCdx2 and iGata6 lines, then picked a clonal line that strongly induces Oct4 on addition of dox.

I set up parallel timecourses of induced and uninduced samples, transferring them to SL media with or without dox. Cells exposed to dox rapidly upregulated Oct4, while Oct4 levels remained consistent in the uninduced samples (Figure 5.10b). Uninduced cells showed partial downregulation of Nanog, consistent with previously described heterogeneity and reduced expression of Nanog in SL conditions compared to 2iL (Chambers et al., 2007; Sousa et al., 2018; Ying et al., 2008). On the other hand, Oct4 overexpressing cells showed more rapid and complete repression of Nanog (Figure 5.10b).

Figure 5.10

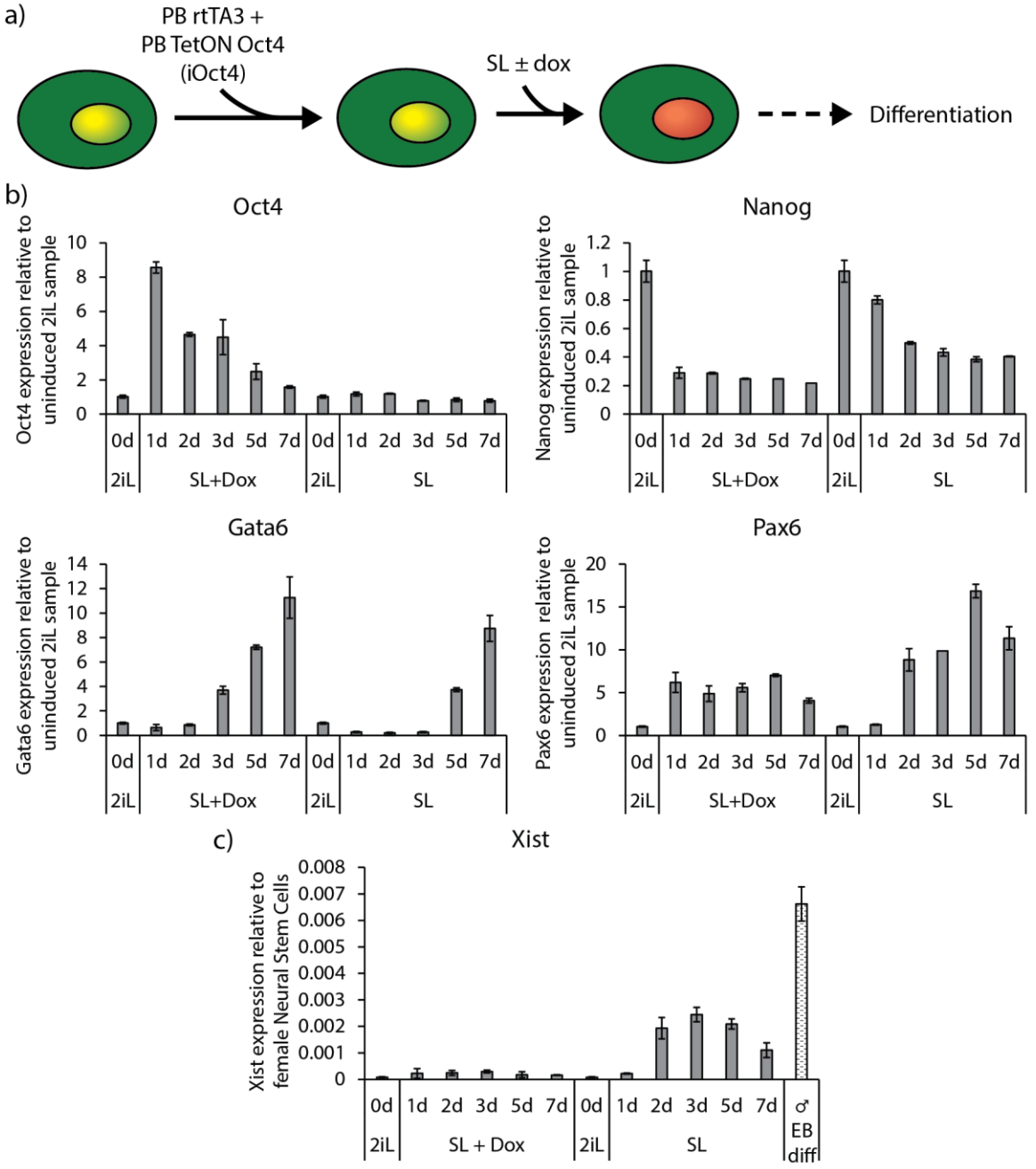



Figure 5.10 Overexpression of Oct4 in male ESCs results in differentiation with repression of Xist
a) Schematic for generating and inducing iOct4 male ESCs. Wild-type male ESCs were transfected with a randomly integrating constitutive rtTA construct and a randomly integrating TRE Oct4 cassette (together iOct4). b) Replicate wells of iOct4 ESCs were maintained in N2B27-2iL and transferred to differentiation-permissive SL, with or without addition of dox. Expression of Oct4, the pluripotency factor Nanog, the ectoderm marker Pax6 and the mesoderm marker Gata6 were detected by RT-qPCR. Values represent the mean of three replicate qPCR reactions and error bars represent ± SD; values were calculated using the ΔΔCt method using the housekeeping gene Gapdh as a reference gene and normalized to 0 hours. c) Samples in b) were examined for expression of Xist which was detected by RT-qPCR. Values represent the mean of three replicate qPCR reactions and error bars represent ± SD; values were calculated by the ΔΔCt method using the housekeeping gene Gapdh as a reference gene and normalized to female NSCs as in Figure 5.7b, with the peak Xist expression detected in male ESCs during embryoid body differentiation shown for scale.

Uninduced cells transferred to these conditions showed a degree of differentiation, with upregulation of the lineage markers Pax6 (ectoderm) and Gata6 (mesoderm) (Figure 5.10b). Concurrently they showed mild upregulation of Xist; note that this also serves to demonstrate that while Xist is upregulated on transition of ESCs to SL, the scale is incomparable to that seen on overexpression of Cdx2 or Gata6 where levels are 100-500 fold greater. Overexpression of Oct4 resulted in faster induction of lineage markers. Despite this, and the far greater percentage of the cells undergoing differentiation, upregulation of Xist was greatly reduced in Oct4-overexpressing cells compared to uninduced cells simply transiting to SL. Not only does this demonstrate that Xist upregulation is not caused by our dox-inducible transgene expression system, it also indicates that Oct4 plays an active role in the repression of Xist even as the rest of the naïve transcription network collapses.

## 5.3. Discussion

Random X chromosome inactivation is a common mechanism by which eutherians achieve dosage compensation of X linked genes. Female cells of the preimplantation epiblast, and ESCs *in vitro*, have two active X chromosomes, and one of these undergoes inactivation at the onset of differentiation. Xist is a master regulator of this process, coating the prospective inactive X chromosome and recruiting chromatin modifiers to inhibit expression. Meanwhile, differentiating male cells show very transient Xist upregulation and rapidly reverse any associated gene inactivation that may occur. As described in the previous chapter, the long-standing model of XCI suggests that a process of counting X chromosomes and choice of the presumptive inactive X follows the decision to initiate differentiation, followed by upregulation of Xist on the chosen chromosome. Work in our group suggests that events unfold differently, with Xist upregulated as a direct consequence of the collapse of the naïve network. Under this new model, counting and choice occur subsequently and result in downregulation of Xist on chromosomes that will not be inactivated.

### 5.3.1. Xist is misregulated in male cells following depletion of Oct4

The observation that male cells upregulate Xist, the principal factor in the control of X chromosome inactivation (XCI), following depletion of Oct4 is somewhat unexpected. Literature suggests that during differentiation female cells specifically upregulate Xist from the prospective inactive X chromosome (Panning et al., 1997; Sheardown et al., 1997), while male cells entirely silence Xist from the low levels observed in ESCs maintained in SL (Panning et al., 1997; Sheardown et al., 1997). Since male cells have a single X chromosome, it is expected that they should not upregulate Xist as counting

precedes the onset of inactivation. However, recent data suggest that upregulation of Xist is connected to the dissolution of the naïve pluripotency network (Sousa et al., 2018). Since this happens rapidly on removal of Oct4, it is not entirely surprising that Xist is upregulated.

However, following genetic ablation of Oct4 I observe a far greater expression of Xist in male cells than is observed in normal differentiation of male ESCs or on loss of Nanog (Sousa et al., 2018), with Xist levels continuing to rise almost in line with female cells (Figure 5.2b). Initially, it is somewhat surprising that Xist is upregulated so highly and persistently in male ESCs on loss of Oct4, since Xist is not expressed at all in male trophoblast cells. However, it is important to note that during development *in vivo*, Xist expression in the extraembryonic lineages is regulated in a different manner than later in embryonic lineages. The maternal Xist allele is silenced through epigenetic imprinting during oogenesis (Fukuda et al., 2014; Inoue et al., 2017). It is reactivated only in the naïve epiblast (Silva et al., 2009), and therefore remains inactive in extraembryonic tissues in males, while paternal Xist expression maintains silencing of the paternal X chromosome in the extraembryonic compartment in females. It has been shown that removal of these imprints in early embryogenesis results in Xist expression from both X chromosomes in female trophectoderm, leading to initiation of XCI on both chromosomes (Inoue et al., 2017). Furthermore, it has been suggested that inheritance of two maternal X chromosomes in parthenogenotes, which results in failure to inactivate either chromosome due to the lack of Xist expression, may be responsible for the very low survival rate shortly after implantation (~25% survive to E9.5) (Tada and Takagi, 1992). While reports vary (Kaufman et al., 1978; Rastan et al., 1980), it appears this could particularly affect extraembryonic tissues, with many implanted parthenogenotes lacking a discernible ectoplacental cone or extraembryonic endoderm (Tada and Takagi, 1992). However, a proportion of cells eventually escape the imprinted inactivation of Xist and successfully inactivate a single X chromosome (Nesterova et al., 2001). It seems plausible that the machinery for counting and inactivating Xist may be absent in less potent cells undergoing transdifferentiation to a trophectoderm-like fate since this differentiation normally occurs within the very early embryo. It may be that following induction of Xist on exit from pluripotency, silencing of the lncRNA cannot occur unless cells are progressing normally towards an epiblast-like identity.

This would explain the continued accumulation of Xist RNA in male cells following depletion of Oct4; since Xist is not imprinted in ESCs, the induced trophoblast-like cells would naturally express all Xist

alleles. It would also predict that the trophoblast-like cells generated by removing Oct4 should be defective in proliferation and survival. I attempted to generate a trophoblast stem cell line following Oct4 depletion in male ESCs; while I was able to serially passage a line that expressed trophectoderm markers Elf5 and Cdx2 (Figure 5.3a and b), cells were very slow growing and a large percentage of the population was lost every passage. After three passages Xist was not detectable (Figure 5.3b), perhaps suggesting that a small percentage of cells are able to recover their normal X status as in parthenogenotes, and these may have a selective advantage.

### 5.3.2. Xist expression is a common feature of extraembryonic differentiation from ESCs

Depletion of Oct4 results in trophoblast-like differentiation. As described above, this results in strong and persistent Xist upregulation in male cells. I considered this to be evidence that collapse of the pluripotency network was the proximal cause of Xist induction and hypothesized that extraembryonic cells lack the machinery for either counting, choice or Xist repression.

I therefore sought to confirm that alternative methods of inducing extraembryonic differentiation of male ESCs would result in stable upregulation of Xist. To do so, I established cells lines for inducible expression of Cdx2 or Gata6, which have previously been demonstrated to stimulate trophoblast and extraembryonic endoderm differentiation of ESCs respectively. After demonstrating appropriate transgene induction and subsequent expression of relevant extraembryonic markers, I proceeded to analyse Xist expression.

I have found that male cells exiting pluripotency and differentiating towards trophectoderm show persistent upregulation of Xist to comparable levels to female somatic cells. This is the case whether Oct4 depletion or Cdx2 induction are used to drive the change in cell identity (Figure 5.7b and Figure 5.8). Notably, I did observe rapid downregulation of Oct4 following overexpression of Cdx2, suggesting that the upregulation of Xist in these two systems could occur via a common mechanism. However, it is also possible that Cdx2 causes Xist to be expressed independent of Oct4 levels. Ongoing work to test this will utilize a fixed-Oct4 ESC background in which Oct4 expression is constitutively maintained. Therefore, if upregulation of Xist is blocked on overexpression of Cdx2, this will strongly suggest that the downregulation of Oct4 in normal cells is vital for this phenomenon to occur, demonstrating that Oct4 is a key negative regulator of Xist.

Furthermore, the transition of ESCs to an alternative extraembryonic identity, induction of Gata6 leading to extraembryonic endoderm differentiation, was also marked by significant upregulation of Xist (Figure 5.7b and Figure 5.8). Together with published observations that reduced Nanog expression correlates with enhanced Xist expression in naïve ESCs, and that Xist is rapidly upregulated in male cells following downregulation of the naïve transcriptional network (Sousa et al., 2018), this provides strong evidence that pluripotency-associated factors play a major role in the repression of Xist.

Extraembryonic tissues regulate dosage compensation through the imprinted inactivation of Xist on the maternal X chromosome. They are expected to express Xist from any non-imprinted X inactivation centre, a hypothesis that is confirmed by experiments studying parthenogenetic embryos. Together, these factors explain why Xist upregulation is so dramatic and protracted in my experiments inducing extraembryonic differentiation from ESCs.

What remains unclear, however, is whether these cells continue to express the machinery to enact XCI and what the outcome of this would be. This system could be useful for further studies into the mechanism of X chromosome counting, choice and Xist inactivation, as they appear to be defective in at least one of these stages as these male cells fail to downregulate Xist as in somatic differentiation, instead maintaining a high level of expression over at least several days.

### 5.3.3. Oct4 overexpression prevents Xist upregulation

Remarkably, it appears that increased Oct4 expression is sufficient to prevent upregulation of Xist during somatic differentiation from male ESCs (Figure 5.10c). It should be noted that the level of Xist upregulation in this system in the absence of Oct4 is comparable to very early female differentiation, after which female cells rapidly induce Xist to far greater levels whereas Xist subsequently becomes silenced in male cells. It would therefore be interesting to repeat the Oct4 overexpression induced differentiation experiments in female cells to determine whether Oct4 is capable of repressing Xist from far stronger inductive influences. It is also important to keep in mind that Oct4 levels remain relatively consistent for some time in normal differentiation; wild-type levels of Oct4 are clearly not sufficient to prevent Xist induction. However, it has previously been shown that Oct4 binds to regulatory elements of Tsix and Tsx, inhibitors of Xist, and promotes their expression (Donohoe et al., 2009).

It therefore appears that Xist is upregulated at the onset of both somatic and extraembryonic lineages from male ESCs. This appears to be caused by the general dismantling of the naïve transcription factor network, but can be prevented through overexpression of Oct4. In somatic differentiation, counting and choice result in the silencing of Xist from X chromosomes that are not destined to undergo XCI. Meanwhile, during extraembryonic differentiation the process of counting, choice, or Xist silencing fails, and Xist continues to be expressed to high levels even in male cells. Together this demonstrates novel differences in the machinery associated with XCI between somatic and extraembryonic differentiation; this could provide a platform for future work to establish the exact factors and mechanisms involved in this process. Additionally, it demonstrates a role for Oct4 in preventing Xist induction. Speculatively, this could be related to the mechanism by which Xist expression is silenced following counting and choice, and I hope that future work will elucidate whether this is the case.

# Chapter 6: General Discussion

### 6.1.1. Transcriptional regulation by Oct4

Typically Oct4 is viewed as a core pluripotency factor, and a master regulator of the naïve state. It is placed as one of a trifecta of key factors, together with Sox2 and Nanog, which together stabilize naïve pluripotency. The three factors were identified as essential for the maintenance of pluripotency, while other factors were found to be redundant in their activity. However, over time it has become increasingly clear that this is not the case. In the absence of external differentiation stimuli, Nanog is not essential for the maintenance of pluripotency in mouse ESCs. Meanwhile, overexpression of Oct4 is sufficient to render Sox2 dispensable for their self-renewal and differentiation. It appears that this leaves Oct4 at the pinnacle of a hierarchy of pluripotency factors.

In a sense, this may be true; from the work presented here, it appears that the binding of Oct4 may be essential for the expression of important pluripotency genes (Figure 4.7). This requirement for Oct4 likely underscores the importance of this factor in reprogramming, where activation of endogenous pluripotency loci is a key step. The rapid reduction in enhancer RNA expression (Figure 4.15) and histone modifications associated with active enhancers (Figure 4.14) suggests that Oct4 is required for the licensing of these elements. It has also been shown that the presence of a promotor is required for expression of enhancer RNAs (Kim et al., 2010); this suggests that a possible essential activity of Oct4 is to induce enhancer-promoter interactions (Figure 6.1), and that this DNA looping could be lost in the absence of Oct4.

Figure 6.1

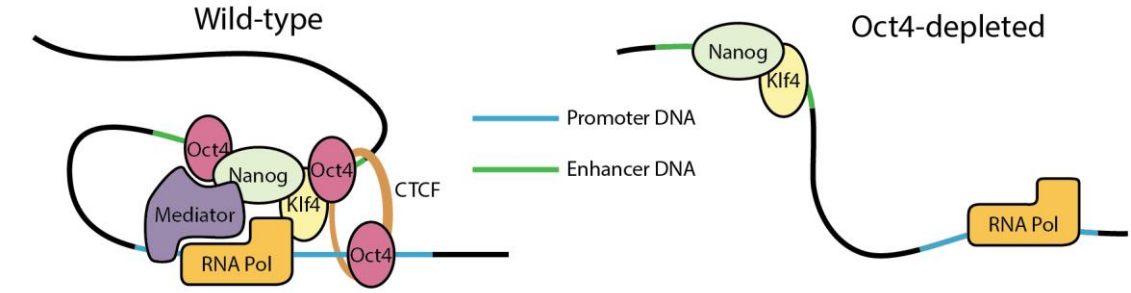



Figure 6.1 A potential role for Oct4 in establishing enhancer-promoter interactions
Schematic showing how Oct4 could be required for enhancer-promoter interactions by recruiting Mediator and CTCF (left). In the absence of Oct4 this looping would be lost, preventing transactivation of gene expression by enhancer-bound factors (right).

It is interesting to note that binding of CTCF and the Mediator complex, both of which support long-range enhancer-promoter interactions (Donohoe et al., 2009; Kagey et al., 2010), is overrepresented at Oct4 binding sites (Göke et al., 2011), and Mediator strongly co-localizes with Oct4-Nanog-Sox2 triple bound loci (Göke et al., 2011). Furthermore, it has been demonstrated that Oct4 physically interacts with CTCF (Donohoe et al., 2009). Moving forward, it would be interesting to examine how chromatin architecture is modified upon loss of Oct4 through chromatin conformation capture studies, in order to identify whether promotion of enhancer looping is a key property of Oct4 in maintaining expression of key pluripotency factors.

However, the role of Oct4 is complex, and does not simply involve activation of pluripotency-associated genes. It has previously been shown that a constitutive trans-activating Oct4 fusion protein can maintain ESC self-renewal; however, these cells are defective in differentiation, demonstrating that a repressive role of Oct4 is essential for its normal function in ESCs (Hammachi et al., 2012). This is in keeping with the fact that overexpression of wild-type Oct4 induces differentiation of ESCs, while a reduced level of Oct4 enhances self-renewal while blocking differentiation. Mechanistically, this may be linked to the enhanced binding of Nanog to the genome following reduction of Oct4 levels that I have observed (Figure 4.9c, Figure 4.10b, Figure 4.11, Figure 4.12). It may be that overexpression of Oct4 leads to sequestration of pluripotency factors at non-essential enhancer elements, depleting them from key loci, or even the formation of complexes in the nucleoplasm without binding to chromatin at all (Figure 6.2). Alternatively, it may be that at higher concentrations enhancer elements spend a greater proportion of time with Oct4 bound, resulting in physical occlusion of binding sites for other factors or allowing sufficient time for the many inhibitory chromatin modifiers that interact with Oct4 to decommission active enhancers (Figure 6.2).

Figure 6.2

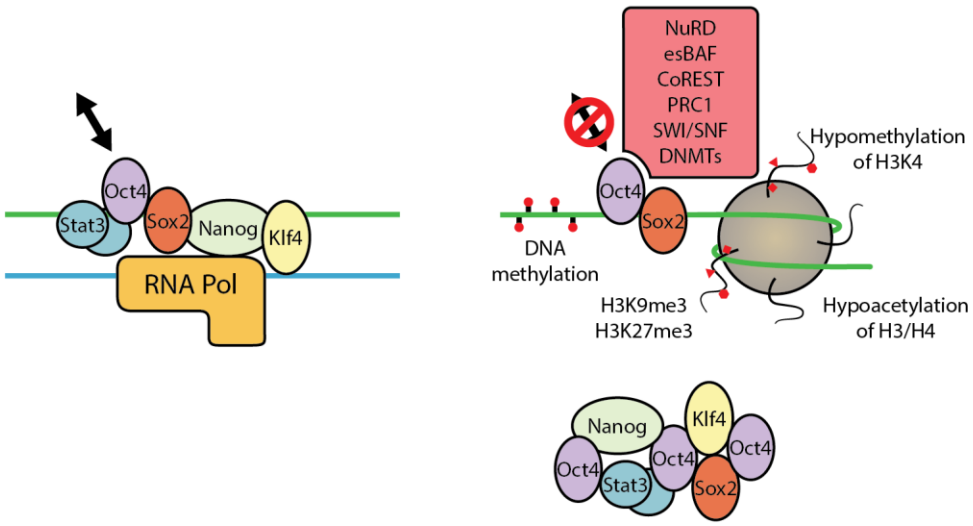



Figure 6.2 Models of mechanisms by which Oct4 overexpression may induce repression
When Oct4 is lowly expressed it may transiently interact with chromatin; meanwhile other pluripotency factors robustly interact with enhancer elements to promote expression of genes that enhance self-renewal (left). As Oct4 levels increase (right), Oct4 occupancy of enhancer elements may increase, resulting in greater recruitment of repressive complexes leading to epigenetic silencing of self-renewal genes. Alternatively, excess Oct4 may bind to and sequester pluripotency factors in the absence of chromatin.

In order to investigate the mechanism by which overexpression of Oct4 leads to collapse of the naïve network in more detail, there are several routes that could be pursued. Exploring the genome-wide binding of various transcription factors in wild-type and Oct4-overexpressing ESCs would give an indication of whether Oct4 is recruiting pluripotency factors away from key loci and spreading them across the genome, while quantitative mass spectrometry of non-chromatin-bound Oct4 binding partners would indicate whether Oct4 is sequestering proteins within the nucleoplasm.

It would also be interesting to investigate the recruitment of chromatin modifying complexes by Oct4. While protein interactions have been routinely observed between Oct4 and members of various repressive complexes, it is unclear how positive regulators are recruited to some enhancer elements and negative regulators to others. It is unknown whether Oct4 residence time has a role in this decision; this could be examined through single-molecule tracking of Oct4 combined with labelling of activatory loci such as the FGF4 enhancer and repressive loci such as Cdx2. If there appears to be a significant difference between such loci, this could be further tested by introducing point mutations to create stronger or weaker Oct4 binding sites, and therefore alter the residence time of Oct4.

However, it may be that the chromatin context of the locus or co-occupancy by other transcription factors that effects this difference in function. This is harder to test since, by their nature, repressive

loci will be bound by repressive factors and positive regulatory elements will be bound by trans-activating factors. It is possible to test whether transiently altering the chromatin environment at repressed loci, through recruitment of P300 by a dCas9 (catalytically inactive Cas9, an RNA-mediated DNA binding protein) fusion protein for instance, alters the composition of bound factors, and whether a repressive chromatin environment is reinstated or if this permanently alters the function of the regulatory element. More thorough profiling of the factors bound at positive and negative regulatory elements could also be informative. This could be examined through enChIP (engineered DNA-binding-molecule-mediated ChIP), which utilizes custom DNA-binding factors such as dCas9 to precipitate specific loci of interest, after which the bound proteins can be identified though sensitive mass spectrometry. However, this technique is still being developed and therefore costly and inefficient, and it is unclear whether it provides sufficient sensitivity for identifying novel factors bound at complex regulatory elements, particularly in a repressive chromatin context.

### 6.1.2. Oct4 as a reprogramming factor

Oct4 is known to be an important factor in the process of somatic cell reprogramming, as well as in transdifferentiation. However, it is common for Oct4 to be used at levels never normally seen in a biological context. This leads to questions as to whether the expression of Oct4 seen in the early embryo is sufficient to induce cell fate transitions or if this is a neomorphic effect of extreme overexpression of Oct4. It is well known that the precise level of Oct4 has a profound impact on its ability to influence cell fate decisions. Previous work has indicated that low levels of Oct4 restrict the ability of naïve cells to exit pluripotency (Karwacki-Neisius et al., 2013; Radzisheuskaya et al., 2013). Here we observe that an ESC level of Oct4 is permissive for the fate transition from somatic cells back to the naïve state (Figure 3.4, Figure 3.5, Figure 3.6). This is supported by the recent finding that activation of the endogenous Oct4 locus using an artificial transcription factor is sufficient to induce somatic cell reprogramming (P. Liu et al., 2018); however, in that work the level of Oct4 expressed during reprogramming in single cells that successfully generate iPSCs remains unclear. Additionally, preliminary work being performed in our group suggest that this level of Oct4 may facilitate the transdifferentiation of somatic cells, as well as reprogramming of E12.5-derived MEFs to pluripotency. Looking to the future, it would be interesting to explore whether different levels of Oct4 facilitate different stages of the reprogramming process. It is interesting that in the course of reprogramming $Oct4^{-/-}$ cells with transgenic Oct4 there is some form of restriction such that all successful

reprogramming events occur in cells with an ESC level of Oct4 (Radzisheuskaya et al., 2013). It may be that higher levels are deleterious to the expression of pluripotency factors as seen in stable naïve cells. With this in mind, it would be informative to examine the reprogramming efficiency of somatic cells with different defined levels of Oct4. This could perhaps be achieved by targeted knock-in of Oct4 with enhancer elements of differing strengths to the Rosa26 locus in an $Oct4^{-/-}$ background. Alternatively, the kinetics of Oct4 expression over the course of reprogramming could be tracked in single cells alongside a series of markers indicative of a productive reprogramming trajectory (Schwarz et al., 2018; Zunder et al., 2015). Tagging Oct4 with the short fragment of a split fluorescent protein (Cabantous et al., 2005; Feng et al., 2017) could permit accurate tracking of Oct4 levels in real time in live cells. In this system, a large fragment of GFP that is not significantly fluorescent is expressed constitutively, while a short, complementary fragment is tagged to the protein of interest. This tag can be as short as 17 amino acids. When co-expressed, the fragments self-associate and form a functional fluorophore, without requiring lengthy fluorophore folding times or causing significant disruption to the activity of the tagged protein. This would make it possible to track the level of Oct4 expressed over time within cells, and retrospectively identify those that successfully reprogram, in order to find out the level of Oct4 that is most beneficial at different stages in the reprogramming context.

### 6.1.3. Oct4 in pluripotency and differentiation

Nevertheless, the fact that a wild-type level of Oct4 is permissive both for differentiation and reprogramming implies that rather than being a specifier of the naïve state, Oct4 may act primarily as an architect of pluripotency in a broader sense, holding the key to cell fate transitions when present in an appropriate quantity. It may be that the secret to maintaining the naïve state lies less in the specifics of the transcription factor network and more in the signalling conditions of the preimplantation epiblast, with Oct4 allowing cells to progress forward as soon as the embryo implants and the environment shifts. It remains unclear what signals within the embryo maintain and dissolve the naïve state. *In vitro*, autocrine FGF4 signalling is a major driver of differentiation; cells of the naïve epiblast robustly express FGFR1 and secrete FGF4, so it is unclear how they retain their pluripotent identity (Kang et al., 2017; Molotkov et al., 2017). Additionally, LIF signalling is not required to sustain the naïve epiblast (Stewart et al., 1992; Ware et al., 1995; Yoshida et al., 1996) unless embryos enter into diapause (Nichols et al., 2001). It seems likely then that BMP, WNT and Activin/Nodal signalling may play a more important role in the maintenance of naïve pluripotency *in vivo*, but the precise

requirement of each of these remains unclear (Menchero et al., 2017). The shift in identity from naïve to primed, and the concurrent changes in morphology and cell-cell interactions, are presumably driven by signals from the trophectoderm, but this too is poorly understood. It is interesting to note that Oct4 continues to be expressed in PGCs, and could therefore play a key role in their capacity to convert to EGCs in appropriate conditions. The role of Oct4 in these other contexts is poorly understood, and future work in new systems for studying early embryonic differentiation *in vitro*, such as gastruloids (van den Brink et al., 2014) and ETX embryo structures (Sozen et al., 2018), may play a key role in uncovering the function of Oct4 as cells continue forwards in development.

Recent work has uncovered conditions for maintaining human pluripotent cells in a naïve-like state (Takashima et al., 2014; Theunissen et al., 2014). While it remains unclear precisely how these relate to the naïve epiblast of human embryos (Bates and Silva, 2017), they provide the best model system to date for studying the earliest fate transitions following epiblast specification in humans.

Interestingly, differentiation of these cells is a slow and drawn-out process, requiring a stage of priming prior to efficient exit from pluripotency (Guo et al., 2016; Takashima et al., 2014), mirroring the far longer process of implantation, polarization and priming observed in non-human primate embryos *in vivo* (Nakamura et al., 2016). This work has led to the suggestion that cells may pass through a separate identity, perhaps equivalent to mouse EpiLCs, during their transition to the primed state (Smith, 2017). Using this knowledge, work is ongoing to establish conditions to maintain this state *in vitro* (Kinoshita and Smith, 2018; Mulas et al., 2017), particularly using mouse cells. This has been labelled the formative state, as it has been suggested that it is during this transient phase that cells acquire the capacity for somatic differentiation (Smith, 2017). This is based on the observations that immediately following exit from naïve-specific growth conditions cells are unresponsive to differentiation stimuli (Mulas et al., 2017) and that the dissolution of the naïve transcription network precedes induction of lineage markers (Kalkan et al., 2017; Kalkan and Smith, 2014).

It is particularly interesting to consider the phenotype of Oct4 knockout in naïve ESCs in this context. We have seen that depletion of Oct4 results in the dismantling of the naïve transcriptional network in the absence of forward differentiation (Figure 4.8). This is eventually followed by upregulation of trophectoderm-associated gene expression, as readily observed following rapid ablation of Oct4 at the protein level. It is therefore tempting to speculate that Oct4 may be required for entry to the formative state or for gaining the capacity to undergo somatic differentiation and that consequently Oct4-null

cells must pass to an identity that does not require transit through the formative state. Research into the role of Oct4 in the formative state may well bring rapid advances to our understanding of how pluripotent cells gain the capability to undergo multi-lineage differentiation and form all the cell types of the embryo proper. Evidently this fits well with the role of Oct4 established throughout this body of work as primarily a facilitator of cell fate transitions and identity changes rather than specifically acting to maintain pluripotency as previously assumed.

### 6.1.4. Control of Xist repression by Oct4 and in extraembryonic differentiation

Recent work from our group has suggested that Xist is not expressed in ground state naïve cells maintained in 2iL conditions, but it upregulated from all alleles as cells exit pluripotency (Sousa et al., 2018). Following counting and choice, Xist expression is subsequently restricted to the chromosomes that will undergo complete XCI. Data presented in this work supports this hypothesis; the fact that Xist upregulation is observed on depletion of Oct4 (Figure 5.2), despite the fact that these cells do not upregulate markers of primed pluripotency (Figure 4.8), suggests that it is loss of naïve-specific factors that triggers this event rather than progression to a more advanced stage of pluripotency. However, this leads to the question of what causes the subsequent downregulation of Xist on chromosomes that are not destined to undergo XCI.

Mutation of Tsix in males results in initiation of XCI in a small proportion of cells (Luikenhuis et al., 2001; Sado et al., 2002; Vigneau et al., 2006), while females always choose to silence the mutant chromosome (Lee and Lu, 1999; Luikenhuis et al., 2001). Recent work suggests that a small proportion of female Tsix mutants actually undergo XCI from either the mutant allele or both alleles and the apparent non-random inactivation is due to the senescence and death of cells that inactivate both chromosomes (Gayen et al., 2015). Additionally, following introduction of a constitutively active promoter to one Tsix allele in females, either the wild-type chromosome is selected and undergoes normal XCI or else both Xist alleles become silenced and XCI does not occur (Stavropoulos et al., 2001). Altogether, this suggests that Tsix plays a role in the silencing of Xist following its initial upregulation at the onset of differentiation.

I propose that Oct4 promotes expression of Tsx; Tsx and Oct4 bound to Xite both co-activate Tsix transcription (Figure 6.3). In the naïve state, binding of Oct4 and other pluripotency factors to intron 1 of Xist, together with this Tsix transcription act to silence Xist. On differentiation, pluripotency factors, perhaps including Oct4, are removed from Xist intron 1 leading to depression and initial

upregulation of Xist. However, Oct4 remains bound at the Tsix and Tsx regulatory elements and their expression increases, leading to silencing of Xist (Figure 6.3). Thus, male cells initially upregulate Xist, but subsequently repress it and do not complete XCI.

Figure 6.3

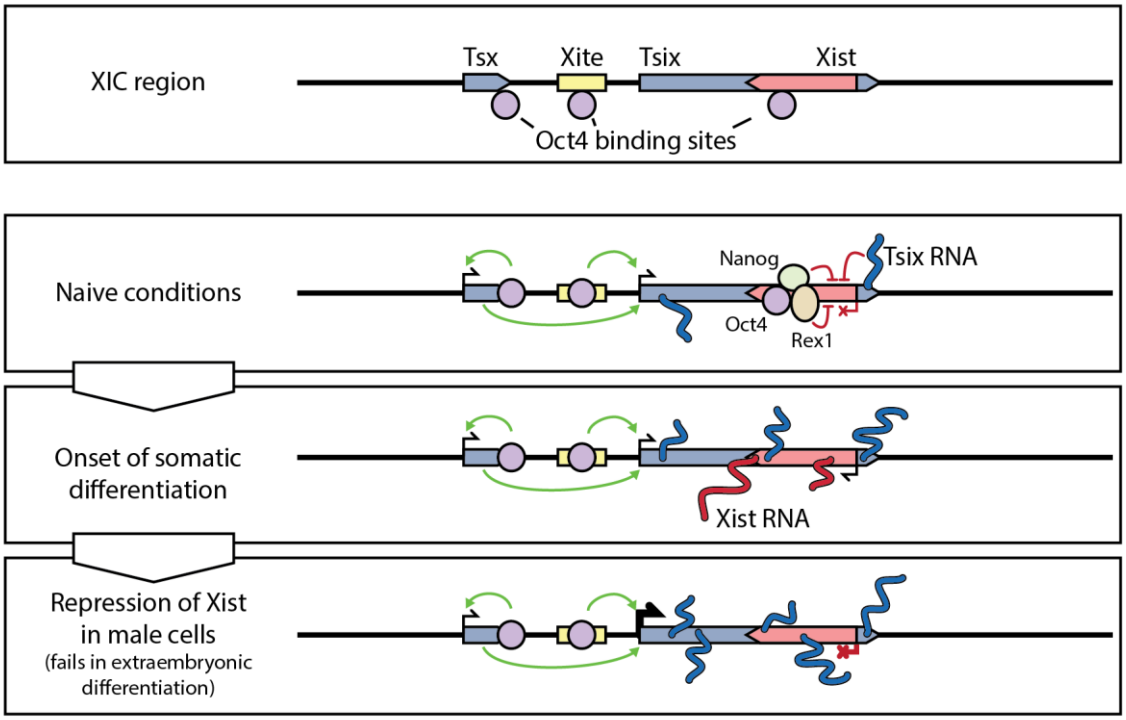



Figure 6.3 A model for regulation of the XIC during male ESC differentiation
Top: Map of the central XIC region. Tsx and Tsix are shown in blue and Xist is shown in red, with arrowheads indicating their direction of transcription. The Xite enhancer element is shown in yellow. Oct4 binding sites are represented by purple circles. Below: In naïve conditions, Oct4 promotes expression of Tsx and induces transcription of Tsix. Tsix RNA (dark blue), together with pluripotency factors bound to Xist intron 1, repress Xist and prevent its expression. At the onset of differentiation, pluripotency factors are downregulated and are no longer bound to Xist intron 1. Oct4 may also be removed. This leads to depression of Xist and the detected Xist RNA (red) in male cells. Meanwhile, Tsx and Tsix expression also increase, possibly due to depression from the Xist intron 1 regulatory element. Continued Oct4 binding to the Tsx and Tsix regulatory elements allows them to be robustly expressed and they result in downregulation and silencing of Xist. The Xist promoter is permanently decommissioned before Oct4 levels decrease significantly.

In the case of Oct4 deletion, expression of Tsx and Tsix would drop, while collapse of the naïve network would initiate upregulation of Xist. In the absence of negative regulators, Xist expression would continue to rise to somatic levels. Rapid repression of Oct4 may be the primary factor that induces prolonged Xist expression in Cdx2- and Gata6-induced extraembryonic differentiation. Alternatively, it could be that other components that promote the silencing of Xist are not present due to the unconventional trajectory of differentiation in these cells.

Figure 6.4

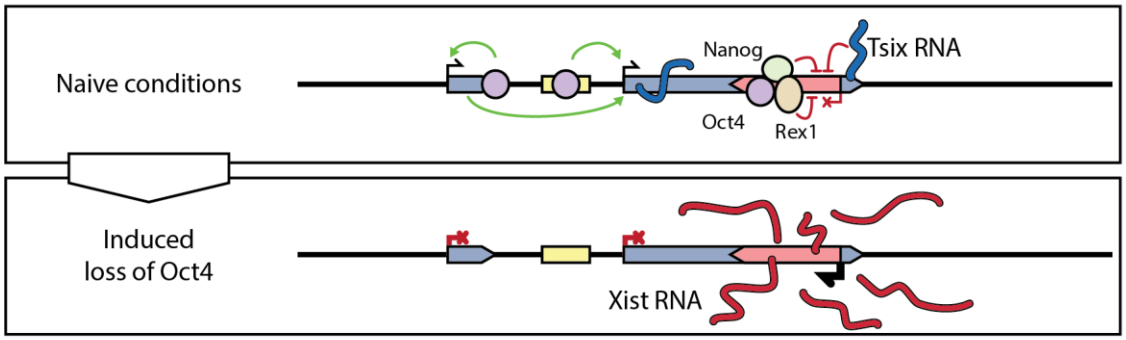



Figure 6.4 A model for the dysregulation of Xist on deletion of Oct4
As in Figure 6.3, Oct4-induced Tsix RNA along with pluripotency factors repress Xist in the naïve state. On depletion of Oct4, however, expression of Tsx and Tsix are compromised, while exit from pluripotency removes repressive factors from Xist intron 1. The combined effect of this is unrestrained upregulation of Xist, allowing it to reach very high and sustained levels of expression.

Regulation in female cells is more complicated, as one chromosome needs to be selected for continued expression of Xist while the other allele is silenced; as this work focused mainly in male cells, more speculation is required to consider how Oct4 may fit in to this process. Introduction of multiple copies of small transgenes containing Xist is required to induce XCI in males, while single copies fail to do so (Herzing et al., 1997). However, heterozygous disruption of Xist in females does not prevent XCI, which still occurs on the chromosome carrying a wild type Xist allele (Marahrens et al., 1998; Penny et al., 1996). This implies that Xist is weakly able to induce counting once it is upregulated, but is not necessary for the process. Other factors must permit the continued expression of Xist in heterozygous females. A recent theory on the mechanism of counting relies on the observation that the both X chromosomes physically pair around the region of the XIC prior to XCI (Augui et al., 2007; Bacher et al., 2006; Xu et al., 2006). The region around the Tsix transcriptional start site and the Xite region (Bacher et al., 2006; Xu et al., 2006), and a region containing the Slc16a2 gene (Augui et al., 2007) all appear to play key roles in this pairing process. It has also been suggested that the reason why multi-copy XIC transgenes can induce XCI in males but single-copy insertions do not relates to the fact that they lack the Slc16a2 locus and therefore display weaker pairing, requiring multiple Tsix/Xite regions to efficiently bring the two chromosomes together (Augui et al., 2007; Xu et al., 2006). Oct4 plays a key role in this process; it has been observed that X-chromosome pairing does not occur following knockdown of Oct4 in female cells (Donohoe et al., 2009). I hypothesize that in the absence of pairing between two functional XIC regions Xist is silenced *in cis* by Tsix, the expression of which is dependent on Oct4. If pairing does occur, one Xist allele is protected from inactivation. This could be through the establishment of an unknown protective mark that prevents silencing of Xist on that chromosome,

or it could simply be through inactivation of Tsix on one chromosome. Once this is established the chromosomes can separate. The unmarked chromosome can undergo silencing as occurs in the absence of pairing. Meanwhile the prospective inactive X upregulates Xist, which coats the chromosome *in cis* and leads to widespread silencing.

Figure 6.5

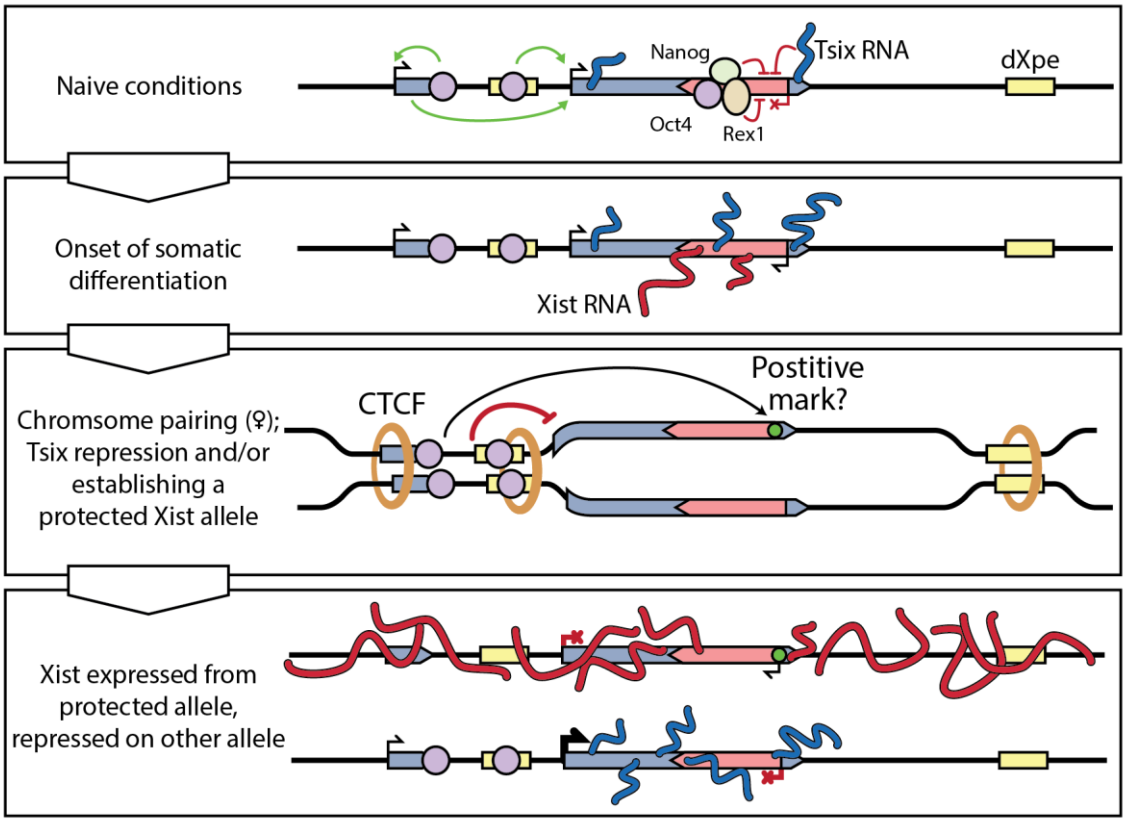



Figure 6.5 A speculative model of the mechanism of XCI
As in Figure 6.3, Oct4 induces Tsx and together they induce expression of Tsix. Tsix RNA and pluripotency factors bound to Xist intron 1 repress Xist in naïve conditions on both chromosomes. Note that ~200kb downstream of Tsix is the distal X chromosome pairing element (dXpe), located in the region of the Slc16a2 gene. At the onset of differentiation, depletion of pluripotency factors leads to upregulation of Xist from both chromosomes, and Tsix and Tsx expression also increase. In the presence of Oct4, CTCF induces pairing of the X chromosomes at the Tsx, Xite and dXpe regions. This either leads to the establishment of a positive mark, or repression of Tsix, on a single chromosome; this is equivalent to 'counting' and 'choice' simultaneously. The chromosomes unpair, and the active X chromosome undergoes normal Xist silencing as in male cells. Xist expression continues to increase from the prospective inactive X chromosome, coating the XIC *in cis*. This domain will continue to expand and recruit repressive factors, leading to silencing of the entire chromosome.

Several factors fall in favour of this hypothesis. Introduction of transgenic pairing regions into autosomes in females reduces the rate of pairing between X chromosomes, and this is associated with failure to induce XCI (Augui et al., 2007; Lee, 2005). In this case, since neither X chromosome is paired to a functional XIC, they would each undergo the Oct4/Tsix-induced Xist silencing as seen in wild-type male cells. Heterozygous deletion of the Tsix start site leads to non-random XCI in the majority

of cells; this would be a combined effect of poor chromosome pairing leading to silencing of Xist on the wild-type allele, and unrestricted Xist upregulation of the mutant allele. Interestingly, homozygous Tsix mutants do undergo random XCI at a greatly reduced efficiency (Lee, 2002); the large proportion of cells display two Xist clouds, but many cells only have one (Lee, 2005). This implies that weaker pairing may still occur and additional elements may contribute to a lesser effect to the silencing of Xist. This could also contribute to the variable penetrance of XCI induction in male cells carrying Tsix mutations.

More work will be required in order to test some of these models. In the simpler case of male differentiation and repression of Xist, it will be important to confirm that Xist upregulation is a prevalent event. This model suggests that removing the Xist intron 1 pluripotency factor binding site and the Tsx/Tsix Oct4 binding sites together should lead to Xist upregulation even within naïve conditions, and should cause unrestricted Xist expression on differentiation in male cells. It may be technically difficult to test this without disrupting the locus, however, though. Another prediction is that restoring Oct4 to Oct4-depleted cells prior to the later stages of XCI might be sufficient to restore Tsx/Tsix expression and silence Xist.

If downregulation of Oct4 is the primary factor causing protracted Xist expression on extraembryonic differentiation, then one would predict that overexpression of Cdx2 in a constitutive Oct4 expression cell line, such as fixed-Oct4 iPSCs, would downregulate Xist as in male somatic differentiation. If this is not the case, then this would be a good model system for investigating other key repressors of Xist. A hypothesis that might be easier to test is that insertion of a constitutive promoter driving Tsix expression could be sufficient to prevent Xist accumulation in male cells on knockout of Oct4 or overexpression of Cdx2.

Testing the model of female XCI remains very difficult. One would expect that homozygous knockout of Tsix would induce sustained Xist expression from both chromosomes, but instead expression appears to be chaotic (Lee, 2002). However, it is unclear whether chromosome pairing is interrupted in this system. It is possible to inhibit transcription with dCas9 which binds to the DNA and blocks the passage of RNA polymerase. This should act without affecting pairing, and could provide a different result, leading to a more complete picture of the role of Tsix in Xist silencing. If an activatory mark is deposited during X chromosome pairing, it is possible that Xist might not be competent to coat the chromosome and induce silencing if chromosome pairing fails to occur; this would result in

continued X-linked expression in male cells expressing Xist. Interestingly, this is what we observe following overexpression of Cdx2 or Gata6.

### 6.1.5. In summary

Altogether, this work demonstrates that Oct4 is a very important and perhaps unique transcription factor. It appears that Oct4 is essential for the expression of several key pluripotency factors in naïve cells, and helps to maintain active epigenetic marks to facilitate this. Despite this, even at physiological levels Oct4 impedes the binding of the pluripotency factor Nanog, a negative impact on the otherwise highly self-sustaining naïve network. This may be a key activity to maintain the vitally important capacity for differentiation in these cells. Interestingly, ablation of Oct4 from these cells seems to entirely prevent forward differentiation, with cells exiting pluripotency entirely and then upregulating extraembryonic genes.

Oct4 also facilitates fate transitions in the opposite direction, from somatic cells back to a naïve identity. It is capable of achieving this at an ESC level, and does not require overexpression to induce reprogramming of *in vivo* differentiated cells.

This work also provides more evidence that differentiation is accompanied by upregulation of Xist from all X chromosomes, and that key steps in the regulation of XCI actually relate to the silencing of Xist from chromosomes that do not require inactivation. Oct4 appears to have a key function in this process as well, contributing to the expression of key regulators that silence Xist when required, and also in establishing inter-chromosomal interactions that lead to the demarcation of one chromosome to be silenced.

### 6.1.6. On the role of Oct4

A great deal of emphasis has been placed on identifying the essential role of Oct4 in the context of naïve pluripotency. From the body of work presented, I hope it is clear that this question is overly simplistic. Oct4 does not play one role, and thus cannot be simply substituted with another gene as is the case for so many naïve factors. Oct4 lies at the heart of many crucial events, including the establishment of the naïve state, maintenance of pluripotency gene expression and competency for differentiation, and control of X chromosome inactivation. In each of these, Oct4 likely acts through many key targets simultaneously, making attempts to rescue knockouts futile. Furthermore, rather than behaving like many other factors and actively enforcing the maintenance of self-renewal and the

naïve identity, Oct4 acts as an arbiter, required for pluripotency while preserving the capacity for differentiation. Rather than a naïve pluripotency factor, I see Oct4 more as a fate transition factor which plays a wider role in development and differentiation *in vivo*, as well as in reprogramming to pluripotency and transdifferentiation *in vitro*. Returning to the metaphor of Waddington's landscape, rather than having a preformed hill, with predefined hills and furrows, it is better to imagine a constantly evolving scene. Stable states are represented by depressions in a plane. Oct4 helps to carve out the indentation representing naïve pluripotency. However, as Oct4 levels increase it reduces the barriers to nearby states and thus encourages identity changes. Other factors such as Nanog or Esrrb simply deepen the recess in which ESCs sit, an effect which is lessened when increasing amounts of Oct4 displace them. While some of the precise mechanistic details of the activity of Oct4 remain to be uncovered, I hope this work ties together many disparate threads of work of many years to form a cohesive explanation of the functions of this complicated factor.

## References

Abel, S., Theologis, A., 1996. Early genes and auxin action. Plant Physiol. 111, 9–17.

Aires, R., Jurberg, A.D., Leal, F., Nóvoa, A., Cohn, M.J., Mallo, M., 2016. Oct4 Is a Key Regulator of Vertebrate Trunk Length Diversity. Dev. Cell 38, 262–74.

Ambrosetti, D.C., Basilico, C., Dailey, L., 1997. Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites. Mol. Cell. Biol. 17, 6321–9.

Ambrosetti, D.C., Schöler, H.R., Dailey, L., Basilico, C., 2000. Modulation of the activity of multiple transcriptional activation domains by the DNA binding domains mediates the synergistic action of Sox2 and Oct-3 on the fibroblast growth factor-4 enhancer. J. Biol. Chem. 275, 23387–97.

Amit, M., Shariki, C., Margulets, V., Itskovitz-Eldor, J., 2004. Feeder layer- and serum-free culture of human embryonic stem cells. Biol. Reprod. 70, 837–45.

Andersson-Rolf, A., Mustata, R.C., Merenda, A., Kim, J., Perera, S., Grego, T., Andrews, K., Tremble, K., Silva, J.C.R., Fink, J., Skarnes, W.C., Koo, B.-K., 2017. One-step generation of conditional and reversible gene knockouts. Nat. Methods 14, 287–289.

Anguera, M.C., Ma, W., Clift, D., Namekawa, S., Kelleher, R.J., Lee, J.T., 2011. Tsx produces a long noncoding RNA and has general functions in the germline, stem cells, and brain. PLoS Genet. 7, e1002248.

Aoki, F., Worrad, D.M., Schultz, R.M., 1997. Regulation of transcriptional activity during the first and second cell cycles in the preimplantation mouse embryo. Dev. Biol. 181, 296–307.

Auclair, G., Guibert, S., Bender, A., Weber, M., 2014. Ontogeny of CpG island methylation and specificity of DNMT3 methyltransferases during embryonic development in the mouse. Genome Biol. 15, 545.

Augui, S., Filion, G.J., Huart, S., Nora, E., Guggiari, M., Maresca, M., Stewart, A.F., Heard, E., 2007. Sensing X chromosome pairs before X inactivation via a novel X-pairing region of the Xic. Science 318, 1632–6.

Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., Lovell-Badge, R., 2003. Multipotent cell

lineages in early mouse development depend on SOX2 function. Genes Dev. 17, 126–140.

Azofeifa, J.G., Allen, M.A., Hendrix, J.R., Read, T., Rubin, J.D., Dowell, R.D., 2018. Enhancer RNA profiling predicts transcription factor activity. Genome Res.

Bacher, C.P., Guggiari, M., Brors, B., Augui, S., Clerc, P., Avner, P., Eils, R., Heard, E., 2006. Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. Nat. Cell Biol. 8, 293–9.

Bao, S., Tang, F., Li, X., Hayashi, K., Gillich, A., Lao, K., Surani, M.A., 2009. Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. Nature 461, 1292–1295.

Bates, L.E., Silva, J.C., 2017. Reprogramming human cells to naïve pluripotency: how close are we? Curr. Opin. Genet. Dev. 46, 58–65.

Beattie, G.M., Lopez, A.D., Bucay, N., Hinton, A., Firpo, M.T., King, C.C., Hayek, A., 2005. Activin A maintains pluripotency of human embryonic stem cells in the absence of feeder layers. Stem Cells 23, 489–95.

Beddington, R.S., Püschel, A.W., Rashbass, P., 1992. Use of chimeras to study gene function in mesodermal tissues during gastrulation and early organogenesis. Ciba Found. Symp. 165, 61-74–7.

Beddington, R.S., Robertson, E.J., 1999. Axis development and early asymmetry in mammals. Cell 96, 195–209.

Bedzhov, I., Zernicka-Goetz, M., 2014. Self-organizing properties of mouse pluripotent cells initiate morphogenesis upon implantation. Cell 156, 1032–44.

Ben-Shushan, E., Thompson, J.R., Gudas, L.J., Bergman, Y., 1998. Rex-1, a gene encoding a transcription factor expressed in the early embryo, is regulated via Oct-3/4 and Oct-6 binding to an octamer site and a novel protein, Rox-1, binding to an adjacent site. Mol. Cell. Biol. 18, 1866–78.

Biddle, J.W., Nguyen, M., Gunawardena, J., 2019. Negative reciprocity, not ordered assembly, underlies the interaction of Sox2 and Oct4 on DNA. Elife 8.

Blaschke, K., Ebata, K.T., Karimi, M.M., Zepeda-Martínez, J.A., Goyal, P., Mahapatra, S., Tam, A., Laird, D.J., Hirst, M., Rao, A., Lorincz, M.C., Ramalho-Santos, M., 2013. Vitamin C induces Tet-

dependent DNA demethylation and a blastocyst-like state in ES cells. Nature 500, 222–6.

Blij, S., Parenti, A., Tabatabai-Yazdi, N., Ralston, A., 2015. Cdx2 efficiently induces trophoblast stem-like cells in naïve, but not primed, pluripotent stem cells. Stem Cells Dev. 24, 1352–65.

Boer, B., Kopp, J., Mallanna, S., Desler, M., Chakravarthy, H., Wilder, P.J., Bernadt, C., Rizzino, A., 2007. Elevating the levels of Sox2 in embryonal carcinoma cells and embryonic stem cells inhibits the expression of Sox2:Oct-3/4 target genes. Nucleic Acids Res. 35, 1773–86.

Boroviak, T., Loos, R., Bertone, P., Smith, A., Nichols, J., 2014. The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. Nat. Cell Biol. 16, 516–28.

Boroviak, T., Loos, R., Lombard, P., Okahara, J., Behr, R., Sasaki, E., Nichols, J., Smith, A., Bertone, P., 2015. Lineage-Specific Profiling Delineates the Emergence and Progression of Naive Pluripotency in Mammalian Embryogenesis. Dev. Cell 35, 366–82.

Borsani, G., Tonlorenzi, R., Simmler, M.C., Dandolo, L., Arnaud, D., Capra, V., Grompe, M., Pizzuti, A., Muzny, D., Lawrence, C., Willard, H.F., Avner, P., Ballabio, A., 1991. Characterization of a murine gene expressed from the inactive X chromosome. Nature 351, 325–9.

Boumil, R.M., Lee, J.T., 2001. Forty years of decoding the silence in X-chromosome inactivation. Hum. Mol. Genet. 10, 2225–32.

Bouniol, C., Nguyen, E., Debey, P., 1995. Endogenous transcription occurs at the 1-cell stage in the mouse embryo. Exp. Cell Res. 218, 57–62.

Bradley, A., Evans, M., Kaufman, M.H., Robertson, E., 1984. Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature 309, 255–256.

Brockdorff, N., Ashworth, A., Kay, G.F., Cooper, P., Smith, S., McCabe, V.M., Norris, D.P., Penny, G.D., Patel, D., Rastan, S., 1991. Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. Nature 351, 329–31.

Brockdorff, N., Ashworth, A., Kay, G.F., McCabe, V.M., Norris, D.P., Cooper, P.J., Swift, S., Rastan, S., 1992. The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. Cell 71, 515–26.

Brons, I.G.M., Smithers, L.E., Trotter, M.W.B., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S.M.,

Howlett, S.K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R.A., Vallier, L., 2007. Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature 448, 191–195.

Brooks, W.S., Helton, E.S., Banerjee, S., Venable, M., Johnson, L., Schoeb, T.R., Kesterson, R.A., Crawford, D.F., 2008. G2E3 is a dual function ubiquitin ligase required for early embryonic development. J. Biol. Chem. 283, 22304–15.

Buckley, S.M., Aranda-Orgilles, B., Strikoudis, A., Apostolou, E., Loizou, E., Moran-Crusio, K., Farnsworth, C.L., Koller, A.A., Dasgupta, R., Silva, J.C., Stadtfeld, M., Hochedlinger, K., Chen, E.I., Aifantis, I., 2012. Regulation of pluripotency and cellular reprogramming by the ubiquitin-proteasome system. Cell Stem Cell 11, 783–98.

Buecker, C., Srinivasan, R., Wu, Z., Calo, E., Acampora, D., Faial, T., Simeone, A., Tan, M., Swigut, T., Wysocka, J., 2014. Reorganization of enhancer patterns in transition from naive to primed pluripotency. Cell Stem Cell 14, 838–53.

Buehr, M., Meek, S., Blair, K., Yang, J., Ure, J., Silva, J., McLay, R., Hall, J., Ying, Q.-L., Smith, A., 2008. Capture of authentic embryonic stem cells from rat blastocysts. Cell 135, 1287–1298.

Burdon, T., Stracey, C., Chambers, I., Nichols, J., Smith, A., 1999. Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells. Dev. Biol. 210, 30–43.

Cabantous, S., Terwilliger, T.C., Waldo, G.S., 2005. Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. Nat. Biotechnol. 23, 102–7.

Cai, L., Friedman, N., Xie, X.S., 2006. Stochastic protein expression in individual cells at the single molecule level. Nature 440, 358–62.

Calaway, J.D., Lenarcic, A.B., Didion, J.P., Wang, J.R., Searle, J.B., McMillan, L., Valdar, W., Pardo-Manuel de Villena, F., 2013. Genetic architecture of skewed X inactivation in the laboratory mouse. PLoS Genet. 9, e1003853.

Canham, M.A., Sharov, A.A., Ko, M.S.H., Brickman, J.M., 2010. Functional heterogeneity of embryonic stem cells revealed through translational amplification of an early endodermal transcript. PLoS Biol. 8, e1000379.

Carey, F.J., Linney, E.A., Pedersen, R.A., 1995. Allocation of epiblast cells to germ layer derivatives during mouse gastrulation as studied with a retroviral vector. Dev. Genet. 17, 29–37.

Chambers, I., 2004. The molecular basis of pluripotency in mouse embryonic stem cells. Cloning Stem Cells 6, 386–91.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., Smith, A., 2003. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell 113, 643–55.

Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., Smith, A., 2007. Nanog safeguards pluripotency and mediates germline development. Nature 450, 1230–1234.

Chen, J., Zhang, Z., Li, L., Chen, B.-C., Revyakin, A., Hajj, B., Legant, W., Dahan, M., Lionnet, T., Betzig, E., Tjian, R., Liu, Z., 2014. Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. Cell 156, 1274–1285.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., Loh, Y.-H., Yeo, H.C., Yeo, Z.X., Narang, V., Govindarajan, K.R., Leong, B., Shahab, A., Ruan, Y., Bourque, G., Sung, W.-K., Clarke, N.D., Wei, C.-L., Ng, H.-H., 2008. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell 133, 1106–17.

Chew, J.-L., Loh, Y.-H., Zhang, W., Chen, X., Tam, W.-L., Yeap, L.-S., Li, P., Ang, Y.-S., Lim, B., Robson, P., Ng, H.-H., 2005. Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. Mol. Cell. Biol. 25, 6031–46.

Cho, O.H., Rivera-Pérez, J.A., Imbalzano, A.N., 2011. Chromatin immunoprecipitation assay for tissue-specific genes using early-stage mouse embryos. J. Vis. Exp.

Choi, H.W., Joo, J.Y., Hong, Y.J., Kim, J.S., Song, H., Lee, J.W., Wu, G., Schöler, H.R., Do, J.T., 2016. Distinct Enhancer Activity of Oct4 in Naive and Primed Mouse Pluripotency. Stem cell reports 7, 911–926.

Chubb, J.R., Trcek, T., Shenoy, S.M., Singer, R.H., 2006. Transcriptional pulsing of a developmental gene. Curr. Biol. 16, 1018–25.

Clemson, C.M., McNeil, J.A., Willard, H.F., Lawrence, J.B., 1996. XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome

structure. J. Cell Biol. 132, 259–75.

Clerc, P., Avner, P., 1998. Role of the region 3' to Xist exon 6 in the counting process of X-chromosome inactivation. Nat. Genet. 19, 249–53.

Coucouvanis, E., Martin, G.R., 1995. Signals for death and survival: a two-step mechanism for cavitation in the vertebrate embryo. Cell 83, 279–87.

Dahéron, L., Opitz, S.L., Zaehres, H., Lensch, M.W., Lensch, W.M., Andrews, P.W., Itskovitz-Eldor, J., Daley, G.Q., 2004. LIF/STAT3 signaling fails to maintain self-renewal of human embryonic stem cells. Stem Cells 22, 770–8.

Dailey, L., Yuan, H., Basilico, C., 1994. Interaction between a novel F9-specific factor and octamer-binding proteins is required for cell-type-restricted activity of the fibroblast growth factor 4 enhancer. Mol. Cell. Biol. 14, 7758–69.

Damjanov, I., Solter, D., Škreb, N., 1971. Teratocarcinogenesis as related to the age of embryos grafted under the kidney capsule. Wilhelm Roux. Arch. Entwickl. Mech. Org. 167, 288–290.

De Santa, F., Barozzi, I., Mietton, F., Ghisletti, S., Polletti, S., Tusi, B.K., Muller, H., Ragoussis, J., Wei, C.-L., Natoli, G., 2010. A large fraction of extragenic RNA pol II transcription sites overlap enhancers. PLoS Biol. 8, e1000384.

Debrand, E., Chureau, C., Arnaud, D., Avner, P., Heard, E., 1999. Functional analysis of the DXPas34 locus, a 3' regulator of Xist expression. Mol. Cell. Biol. 19, 8513–25.

Deshaies, R.J., 1999. SCF and Cullin/Ring H2-based ubiquitin ligases. Annu. Rev. Cell Dev. Biol. 15, 435–67.

DeVeale, B., Brokhman, I., Mohseni, P., Babak, T., Yoon, C., Lin, A., Onishi, K., Tomilin, A., Pevny, L., Zandstra, P.W., Nagy, A., van der Kooy, D., 2013. Oct4 is required ~E7.5 for proliferation in the primitive streak. PLoS Genet. 9, e1003957.

Dharmasiri, N., Dharmasiri, S., Estelle, M., 2005. The F-box protein TIR1 is an auxin receptor. Nature 435, 441–5.

Dietrich, J.-E., Hiiragi, T., 2007. Stochastic patterning in the mouse pre-implantation embryo. Development 134, 4219–31.

Donohoe, M.E., Silva, S.S., Pinter, S.F., Xu, N., Lee, J.T., 2009. The pluripotency factor Oct4 interacts with Ctcf and also controls X-chromosome pairing and counting. Nature 460, 128–32.

Downs, K.M., 2008. Systematic localization of Oct-3/4 to the gastrulating mouse conceptus suggests manifold roles in mammalian development. Dev. Dyn. 237, 464–75.

Ducibella, T., Anderson, E., 1975. Cell shape and membrane changes in the eight-cell mouse embryo: prerequisites for morphogenesis of the blastocyst. Dev. Biol. 47, 45–58.

Ducibella, T., Ukena, T., Karnovsky, M., Anderson, E., 1977. Changes in cell surface and cortical cytoplasmic organization during early embryogenesis in the preimplantation mouse embryo. J. Cell Biol. 74, 153–67.

Esteban, M.A., Wang, T., Qin, B., Yang, J., Qin, D., Cai, J., Li, W., Weng, Z., Chen, J., Ni, S., Chen, K., Li, Y., Liu, X., Xu, J., Zhang, S., Li, F., He, W., Labuda, K., Song, Y., Peterbauer, A., Wolbank, S., Redl, H., Zhong, M., Cai, D., Zeng, L., Pei, D., 2010. Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. Cell Stem Cell 6, 71–9.

Evans, M.J., Kaufman, M.H., 1981. Establishment in culture of pluripotential cells from mouse embryos. Nature 292, 154–156.

Feng, S., Sekine, S., Pessino, V., Li, H., Leonetti, M.D., Huang, B., 2017. Improved split fluorescent proteins for endogenous protein labeling. Nat. Commun. 8, 370.

Ferraris, L., Stewart, A.P., Kang, J., DeSimone, A.M., Gemberling, M., Tantin, D., Fairbrother, W.G., 2011. Combinatorial binding of transcription factors in the pluripotency control regions of the genome. Genome Res. 21, 1055–64.

Festuccia, N., Osorno, R., Halbritter, F., Karwacki-Neisius, V., Navarro, P., Colby, D., Wong, F., Yates, A., Tomlinson, S.R., Chambers, I., 2012. Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. Cell Stem Cell 11, 477–490.

Filipczyk, A., Marr, C., Hastreiter, S., Feigelman, J., Schwarzfischer, M., Hoppe, P.S., Loeffler, D., Kokkaliaris, K.D., Endele, M., Schauberger, B., Hilsenbeck, O., Skylaki, S., Hasenauer, J., Anastassiadis, K., Theis, F.J., Schroeder, T., 2015. Network plasticity of pluripotency transcription factors in embryonic stem cells. Nat. Cell Biol. 17, 1235–46.

Finch, B.W., Ephrussi, B., 1967. Retention of multiple developmental potentialities by cells of a mouse

testicular teratocarcinoma during prolonged culture in vitro and their extinction upon hybridization with cells of permanent lines. Proc. Natl. Acad. Sci. U. S. A. 57, 615–21.

Fong, Y.W., Inouye, C., Yamaguchi, T., Cattoglio, C., Grubisic, I., Tjian, R., 2011. A DNA repair complex functions as an Oct4/Sox2 coactivator in embryonic stem cells. Cell 147, 120–31.

Fowlis, D.J., Ansell, J.D., Micklem, H.S., 1991. Further evidence for the importance of parental source of the Xce allele in X chromosome inactivation. Genet. Res. 58, 63–5.

Fujikura, J., Yamato, E., Yonemura, S., Hosoda, K., Masui, S., Nakao, K., Miyazaki Ji, J., Niwa, H., 2002. Differentiation of embryonic stem cells is induced by GATA factors. Genes Dev. 16, 784–9.

Fukuda, A., Tomikawa, J., Miura, T., Hata, K., Nakabayashi, K., Eggan, K., Akutsu, H., Umezawa, A., 2014. The role of maternal-specific H3K9me3 modification in establishing imprinted X-chromosome inactivation and embryogenesis in mice. Nat. Commun. 5, 5464.

Gagliardi, A., Mullin, N.P., Ying Tan, Z., Colby, D., Kousa, A.I., Halbritter, F., Weiss, J.T., Felker, A., Bezstarosti, K., Favaro, R., Demmers, J., Nicolis, S.K., Tomlinson, S.R., Poot, R. a, Chambers, I., 2013. A direct physical interaction between Nanog and Sox2 regulates embryonic stem cell self-renewal. EMBO J. 32, 2231–47.

Gardner, R.L., 1968. Mouse chimeras obtained by the injection of cells into the blastocyst. Nature 220, 596–7.

Gayen, S., Maclary, E., Buttigieg, E., Hinten, M., Kalantry, S., 2015. A Primary Role for the Tsix lncRNA in Maintaining Random X-Chromosome Inactivation. Cell Rep. 11, 1251–65.

Gillich, A., Bao, S., Grabole, N., Hayashi, K., Trotter, M.W.B., Pasque, V., Magnúsdóttir, E., Surani, M.A., 2012. Epiblast stem cell-based system reveals reprogramming synergy of germline factors. Cell Stem Cell 10, 425–39.

Göke, J., Jung, M., Behrens, S., Chavez, L., O'Keeffe, S., Timmermann, B., Lehrach, H., Adjaye, J., Vingron, M., 2011. Combinatorial binding in human and mouse embryonic stem cells identifies conserved enhancers active in early embryonic development. PLoS Comput. Biol. 7, e1002304.

Golding, I., Paulsson, J., Zawilski, S.M., Cox, E.C., 2005. Real-time kinetics of gene activity in individual bacteria. Cell 123, 1025–36.

Gomes Fernandes, M., Dries, R., Roost, M.S., Semrau, S., de Melo Bernardo, A., Davis, R.P., Ramakrishnan, R., Szuhai, K., Maas, E., Umans, L., Abon Escalona, V., Salvatori, D., Deforce, D., Van Criekinge, W., Huylebroeck, D., Mummery, C., Zwijsen, A., de Sousa Lopes, S.M.C., 2016. BMP-SMAD Signaling Regulates Lineage Priming, but Is Dispensable for Self-Renewal in Mouse Embryonic Stem Cells. Stem cell reports 6, 85–94.

Goto, T., Monk, M., 1998. Regulation of X-chromosome inactivation in development in mice and humans. Microbiol. Mol. Biol. Rev. 62, 362–78.

Goto, Y., Takagi, N., 2000. Maternally inherited X chromosome is not inactivated in mouse blastocysts due to parental imprinting. Chromosome Res. 8, 101–9.

Gray, W.M., del Pozo, J.C., Walker, L., Hobbie, L., Risseeuw, E., Banks, T., Crosby, W.L., Yang, M., Ma, H., Estelle, M., 1999. Identification of an SCF ubiquitin-ligase complex required for auxin response in Arabidopsis thaliana. Genes Dev. 13, 1678–91.

Grumbach, M.M., Morishima, A., Taylor, J.H., 1963. Human sex chromosome abnormalities in relation to dna replication and heterochromatinization. Proc. Natl. Acad. Sci. U. S. A. 49, 581–9.

Guo, G., Huss, M., Tong, G.Q., Wang, C., Li Sun, L., Clarke, N.D., Robson, P., 2010. Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. Dev. Cell 18, 675–85.

Guo, G., Smith, A., 2010. A genome-wide screen in EpiSCs identifies Nr5a nuclear receptors as potent inducers of ground state pluripotency. Development 137, 3185–92.

Guo, G., von Meyenn, F., Santos, F., Chen, Y., Reik, W., Bertone, P., Smith, A., Nichols, J., 2016. Naive Pluripotent Stem Cells Derived Directly from Isolated Cells of the Human Inner Cell Mass. Stem cell reports 6, 437–46.

Guo, G., Yang, J., Nichols, J., Hall, J.S., Eyres, I., Mansfield, W., Smith, A., 2009. Klf4 reverts developmentally programmed restriction of ground state pluripotency. Development 136, 1063–9.

Hall, J., Guo, G., Wray, J., Eyres, I., Nichols, J., Grotewold, L., Morfopoulou, S., Humphreys, P., Mansfield, W., Walker, R., Tomlinson, S., Smith, A., 2009. Oct4 and LIF/Stat3 additively induce Krüppel factors to sustain embryonic stem cell self-renewal. Cell Stem Cell 5, 597–609.

Hamatani, T., Carter, M.G., Sharov, A.A., Ko, M.S.H., 2004. Dynamics of global gene expression changes during mouse preimplantation development. Dev. Cell 6, 117–31.

Hammachi, F., Morrison, G.M., Sharov, A.A., Livigni, A., Narayan, S., Papapetrou, E.P., O'Malley, J., Kaji, K., Ko, M.S.H., Ptashne, M., Brickman, J.M., 2012. Transcriptional Activation by Oct4 Is Sufficient for the Maintenance and Induction of Pluripotency. Cell Rep. 1, 99–109.

Han, D.W., Tapia, N., Joo, J.Y., Greber, B., Araúzo-Bravo, M.J., Bernemann, C., Ko, K., Wu, G., Stehling, M., Do, J.T., Schöler, H.R., 2010. Epiblast stem cell subpopulations represent mouse embryos of distinct pregastrulation stages. Cell 143, 617–27.

Hanel, M.L., Wevrick, R., 2001. Establishment and maintenance of DNA methylation patterns in mouse Ndn: implications for maintenance of imprinting in target genes of the imprinting center. Mol. Cell. Biol. 21, 2384–92.

Hanna, J., Markoulaki, S., Mitalipova, M., Cheng, A.W., Cassady, J.P., Staerk, J., Carey, B.W., Lengner, C.J., Foreman, R., Love, J., Gao, Q., Kim, J., Jaenisch, R., 2009. Metastable pluripotent states in NOD-mouse-derived ESCs. Cell Stem Cell 4, 513–24.

Harris, M.J., Juriloff, D.M., 2007. Mouse mutants with neural tube closure defects and their role in understanding human neural tube defects. Birth Defects Res. A. Clin. Mol. Teratol. 79, 187–210.

Hastreiter, S., Skylaki, S., Loeffler, D., Reimann, A., Hilsenbeck, O., Hoppe, P.S., Coutu, D.L., Kokkaliaris, K.D., Schwarzfischer, M., Anastassiadis, K., Theis, F.J., Schroeder, T., 2018. Inductive and Selective Effects of GSK3 and MEK Inhibition on Nanog Heterogeneity in Embryonic Stem Cells. Stem cell reports 11, 58–69.

Hatano, S.-Y., Tada, M., Kimura, H., Yamaguchi, S., Kono, T., Nakano, T., Suemori, H., Nakatsuji, N., Tada, T., 2005. Pluripotential competence of cells associated with Nanog activity. Mech. Dev. 122, 67–79.

Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., Saitou, M., 2011. Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. Cell 146, 519–32.

Hemmerich, P., Schmiedeberg, L., Diekmann, S., 2011. Dynamic as well as stable protein interactions contribute to genome function and maintenance. Chromosome Res. 19, 131–51.

Herzing, L.B., Romer, J.T., Horn, J.M., Ashworth, A., 1997. Xist has properties of the X-chromosome

inactivation centre. Nature 386, 272–5.

Hirate, Y., Hirahara, S., Inoue, K.-I., Suzuki, A., Alarcon, V.B., Akimoto, K., Hirai, T., Hara, T., Adachi, M., Chida, K., Ohno, S., Marikawa, Y., Nakao, K., Shimono, A., Sasaki, H., 2013. Polarity-dependent distribution of angiomotin localizes Hippo signaling in preimplantation embryos. Curr. Biol. 23, 1181–94.

Hirate, Y., Sasaki, H., 2014. The role of angiomotin phosphorylation in the Hippo pathway during preimplantation mouse development. Tissue barriers 2, e28127.

Hochedlinger, K., Yamada, Y., Beard, C., Jaenisch, R., 2005. Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. Cell 121, 465–77.

Holland, A.J., Fachinetti, D., Han, J.S., Cleveland, D.W., 2012. Inducible, reversible system for the rapid and complete degradation of proteins in mammalian cells. Proc. Natl. Acad. Sci. U. S. A. 109, E3350-7.

Hollnagel, A., Oehlmann, V., Heymer, J., Rüther, U., Nordheim, A., 1999. Id genes are direct targets of bone morphogenetic protein induction in embryonic stem cells. J. Biol. Chem. 274, 19838–45.

Howlett, S.K., Reik, W., 1991. Methylation levels of maternal and paternal genomes during preimplantation development. Development 113, 119–27.

Huang, G., Yan, H., Ye, S., Tong, C., Ying, Q.-L., 2014. STAT3 phosphorylation at tyrosine 705 and serine 727 differentially regulates mouse ESC fates. Stem Cells 32, 1149–60.

Humphrey, R.K., Beattie, G.M., Lopez, A.D., Bucay, N., King, C.C., Firpo, M.T., Rose-John, S., Hayek, A., 2004. Maintenance of pluripotency in human embryonic stem cells is STAT3 independent. Stem Cells 22, 522–30.

Inoue, A., Jiang, L., Lu, F., Zhang, Y., 2017. Genomic imprinting of Xist by maternal H3K27me3. Genes Dev. 31, 1927–1932.

James, D., Levine, A.J., Besser, D., Hemmati-Brivanlou, A., 2005. TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. Development 132, 1273–82.

Jin, J., Liu, J., Chen, C., Liu, Z., Jiang, C., Chu, H., Pan, W., Wang, X., Zhang, L., Li, B., Jiang, C., Ge, X., Xie, X., Wang, P., 2016. The deubiquitinase USP21 maintains the stemness of mouse

embryonic stem cells via stabilization of Nanog. Nat. Commun. 7, 13594.

Joo, J.Y., Choi, H.W., Kim, M.J., Zaehres, H., Tapia, N., Stehling, M., Jung, K.S., Tae, J., Schöler, H.R., Do, J.T., Schöler, H.R., 2014. Establishment of a primed pluripotent epiblast stem cell in FGF4-based conditions. Sci. Rep. 4, 7477.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., Taatjes, D.J., Dekker, J., Young, R.A., 2010. Mediator and cohesin connect gene expression and chromatin architecture. Nature 467, 430–5.

Kalkan, T., Olova, N., Roode, M., Mulas, C., Lee, H.J., Nett, I., Marks, H., Walker, R., Stunnenberg, H.G., Lilley, K.S., Nichols, J., Reik, W., Bertone, P., Smith, A., 2017. Tracking the embryonic stem cell transition from ground state pluripotency. Development 144, 1221–1234.

Kalkan, T., Smith, A., 2014. Mapping the route from naive pluripotency to lineage specification. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 369.

Kalmar, T., Lim, C., Hayward, P., Muñoz-Descalzo, S., Nichols, J., Garcia-Ojalvo, J., Martinez Arias, A., 2009. Regulated fluctuations in nanog expression mediate cell fate decisions in embryonic stem cells. PLoS Biol. 7, e1000149.

Kanatsu, M., Nishikawa, S.I., 1996. In vitro analysis of epiblast tissue potency for hematopoietic cell differentiation. Development 122, 823–30.

Kang, M., Garg, V., Hadjantonakis, A.-K., 2017. Lineage Establishment and Progression within the Inner Cell Mass of the Mouse Blastocyst Requires FGFR1 and FGFR2. Dev. Cell 41, 496–510.e5.

Kang, M., Piliszek, A., Artus, J., Hadjantonakis, A.-K., 2013. FGF4 is required for lineage restriction and salt-and-pepper distribution of primitive endoderm factors but not their initial expression in the mouse. Development 140, 267–79.

Karwacki-Neisius, V., Göke, J., Osorno, R., Halbritter, F., Ng, J.H., Weiße, A.Y., Wong, F.C.K., Gagliardi, A., Mullin, N.P., Festuccia, N., Colby, D., Tomlinson, S.R., Ng, H.-H., Chambers, I., 2013. Reduced Oct4 expression directs a robust pluripotent state with distinct signaling activity and increased enhancer occupancy by Oct4 and Nanog. Cell Stem Cell 12, 531–45.

Katayama, K., Melendez, J., Baumann, J.M., Leslie, J.R., Chauhan, B.K., Nemkul, N., Lang, R.A., Kuan, C.-Y., Zheng, Y., Yoshida, Y., 2011. Loss of RhoA in neural progenitor cells causes the disruption

of adherens junctions and hyperproliferation. Proc. Natl. Acad. Sci. U. S. A. 108, 7607–12.

Kaufman, M.H., Guc-Cubrilo, M., Lyon, M.F., 1978. X chromosome inactivation in diploid parthenogenetic mouse embryos. Nature 271, 547–9.

Kay, G.F., Barton, S.C., Surani, M.A., Rastan, S., 1994. Imprinting and X chromosome counting mechanisms determine Xist expression in early mouse development. Cell 77, 639–50.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., Haussler, D., 2002. The human genome browser at UCSC. Genome Res. 12, 996–1006.

Kepinski, S., Leyser, O., 2005. The Arabidopsis F-box protein TIR1 is an auxin receptor. Nature 435, 446–51.

Kim, J., Chu, J., Shen, X., Wang, J., Orkin, S.H., 2008. An extended transcriptional network for pluripotency of embryonic stem cells. Cell 132, 1049–61.

Kim, J.B., Sebastiano, V., Wu, G., Araúzo-Bravo, M.J., Sasse, P., Gentile, L., Ko, K., Ruau, D., Ehrich, M., van den Boom, D., Meyer, J., Hübner, K., Bernemann, C., Ortmeier, C., Zenke, M., Fleischmann, B.K., Zaehres, H., Schöler, H.R., 2009. Oct4-induced pluripotency in adult neural stem cells. Cell 136, 411–9.

Kim, S.-H., Kim, M.O., Cho, Y.-Y., Yao, K., Kim, D.J., Jeong, C.-H., Yu, D.H., Bae, K.B., Cho, E.J., Jung, S.K., Lee, M.H., Chen, H., Kim, J.Y., Bode, A.M., Dong, Z., 2014. ERK1 phosphorylates Nanog to regulate protein stability and stem cell self-renewal. Stem Cell Res. 13, 1–11.

Kim, T.-K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D.A., Laptewicz, M., Barbara-Haley, K., Kuersten, S., Markenscoff-Papadimitriou, E., Kuhl, D., Bito, H., Worley, P.F., Kreiman, G., Greenberg, M.E., 2010. Widespread transcription at neuronal activity-regulated enhancers. Nature 465, 182–7.

King, H.W., Klose, R.J., 2017. The pioneer factor OCT4 requires the chromatin remodeller BRG1 to support gene regulatory element function in mouse embryonic stem cells. Elife 6.

Kinoshita, M., Shimosato, D., Yamane, M., Niwa, H., 2015. Sox7 is dispensable for primitive endoderm differentiation from mouse ES cells. BMC Dev. Biol. 15, 37.

Kinoshita, M., Smith, A., 2018. Pluripotency Deconstructed. Dev. Growth Differ. 60, 44–52.

Kleinsmith, L.J., Pierce, G.B., 1964. Multipotentiality of single embryonal carcinoma cells. Cancer Res. 24, 1544–51.

Klose, R.J., Yamane, K., Bae, Y., Zhang, D., Erdjument-Bromage, H., Tempst, P., Wong, J., Zhang, Y., 2006. The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. Nature 442, 312–6.

Kojima, Y., Kaufman-Francis, K., Studdert, J.B., Steiner, K.A., Power, M.D., Loebel, D.A.F., Jones, V., Hor, A., de Alencastro, G., Logan, G.J., Teber, E.T., Tam, O.H., Stutz, M.D., Alexander, I.E., Pickett, H.A., Tam, P.P.L., 2014. The transcriptional and functional properties of mouse epiblast stem cells resemble the anterior primitive streak. Cell Stem Cell 14, 107–20.

Kopp, J.L., Ormsbee, B.D., Desler, M., Rizzino, A., 2008. Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells. Stem Cells 26, 903–11.

Kumar, R.M., Cahan, P., Shalek, A.K., Satija, R., DaleyKeyser, Aj., Li, H., Zhang, J., Pardee, K., Gennert, D., Trombetta, J.J., Ferrante, T.C., Regev, A., Daley, G.Q., Collins, J.J., 2014. Deconstructing transcriptional heterogeneity in pluripotent stem cells. Nature 516, 56–61.

Kunath, T., Saba-El-Leil, M.K., Almousailleakh, M., Wray, J., Meloche, S., Smith, A., 2007. FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. Development 134, 2895–902.

Kurimoto, K., Yabuta, Y., Ohinata, Y., Shigeta, M., Yamanaka, K., Saitou, M., 2008. Complex genome-wide transcription dynamics orchestrated by Blimp1 for the specification of the germ cell lineage in mice. Genes Dev. 22, 1617–35.

Kuroda, T., Tada, M., Kubota, H., Kimura, H., Hatano, S., Suemori, H., Nakatsuji, N., Tada, T., 2005. Octamer and Sox elements are required for transcriptional cis regulation of Nanog gene expression. Mol. Cell. Biol. 25, 2475–2485.

Kwon, S.-K., Lee, D.-H., Kim, S.-Y., Park, J.-H., Choi, J., Baek, K.-H., 2017. Ubiquitin-specific protease 21 regulating the K48-linked polyubiquitination of NANOG. Biochem. Biophys. Res. Commun. 482, 1443–1448.

Lam, C.S., Mistri, T.K., Foo, Y.H., Sudhaharan, T., Gan, H.T., Rodda, D., Lim, L.H., Chou, C., Robson, P., Wohland, T., Ahmed, S., 2012. DNA-dependent Oct4-Sox2 interaction and diffusion

properties characteristic of the pluripotent cell state revealed by fluorescence spectroscopy. Biochem. J. 448, 21–33.

Lane, N., Dean, W., Erhardt, S., Hajkova, P., Surani, A., Walter, J., Reik, W., 2003. Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. Genesis 35, 88–93.

Latham, K.E., 1996. X chromosome imprinting and inactivation in the early mammalian embryo. Trends Genet. 12, 134–8.

Lawson, K.A., Pedersen, R.A., 1987. Cell fate, morphogenetic movement and population kinetics of embryonic endoderm at the time of germ layer formation in the mouse. Development 101, 627–52.

Le Bin, G.C., Muñoz-Descalzo, S., Kurowski, A., Leitch, H., Lou, X., Mansfield, W., Etienne-Dumeau, C., Grabole, N., Mulas, C., Niwa, H., Hadjantonakis, A.-K., Nichols, J., 2014. Oct4 is required for lineage priming in the developing inner cell mass of the mouse blastocyst. Development 141, 1001–10.

Lee, J.T., 2005. Regulation of X-chromosome counting by Tsix and Xite sequences. Science 309, 768–71.

Lee, J.T., 2002. Homozygous Tsix mutant mice reveal a sex-ratio distortion and revert to random X-inactivation. Nat. Genet. 32, 195–200.

Lee, J.T., 2000. Disruption of imprinted X inactivation by parent-of-origin effects at Tsix. Cell 103, 17–27.

Lee, J.T., Lu, N., 1999. Targeted mutagenesis of Tsix leads to nonrandom X inactivation. Cell 99, 47–57.

Lee, J.T., Strauss, W.M., Dausman, J.A., Jaenisch, R., 1996. A 450 kb transgene displays properties of the mammalian X-inactivation center. Cell 86, 83–94.

Leitch, H.G., Blair, K., Mansfield, W., Ayetey, H., Humphreys, P., Nichols, J., Surani, M.A., Smith, A., 2010. Embryonic germ cells from mice and rats exhibit properties consistent with a generic pluripotent ground state. Development 137, 2279–87.

Leitch, H.G., McEwen, K.R., Turp, A., Encheva, V., Carroll, T., Grabole, N., Mansfield, W., Nashun,

B., Knezovich, J.G., Smith, A., Surani, M.A., Hajkova, P., 2013. Naive pluripotency is associated with global DNA hypomethylation. Nat. Struct. Mol. Biol. 20, 311–6.

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–60.

Li, M., Yu, J.S.L., Tilgner, K., Ong, S.H., Koike-Yusa, H., Yusa, K., 2018. Genome-wide CRISPR-KO Screen Uncovers mTORC1-Mediated Gsk3 Regulation in Naive Pluripotency Maintenance and Dissolution. Cell Rep. 24, 489–502.

Li, P., Tong, C., Mehrian-Shai, R., Jia, L., Wu, N., Yan, Y., Maxson, R.E., Schulze, E.N., Song, H., Hsieh, C.-L., Pera, M.F., Ying, Q.-L., 2008. Germline competent embryonic stem cells derived from rat blastocysts. Cell 135, 1299–310.

Li, Y., McClintick, J., Zhong, L., Edenberg, H.J., Yoder, M.C., Chan, R.J., 2005. Murine embryonic stem cell differentiation is promoted by SOCS-3 and inhibited by the zinc finger transcription factor Klf4. Blood 105, 635–7.

Li, Y., Zhang, Q., Yin, X., Yang, W., Du, Y., Hou, P., Ge, J., Liu, C., Zhang, W., Zhang, X., Wu, Y., Li, H., Liu, K., Wu, C., Song, Z., Zhao, Y., Shi, Y., Deng, H., 2011. Generation of iPSCs from mouse fibroblasts with a single gene, Oct4, and small molecules. Cell Res. 21, 196–204.

Liang, J., Wan, M., Zhang, Y., Gu, P., Xin, H., Jung, S.Y., Qin, J., Wong, J., Cooney, A.J., Liu, D., Songyang, Z., 2008. Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. Nat. Cell Biol. 10, 731–9.

Liao, Y., Smyth, G.K., Shi, W., 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930.

Lin, Y., Yang, Y., Li, W., Chen, Q., Li, J., Pan, X., Zhou, L., Liu, C., Chen, C., He, J., Cao, H., Yao, H., Zheng, L., Xu, X., Xia, Z., Ren, J., Xiao, L., Li, L., Shen, B., Zhou, H., Wang, Y.-J., 2012. Reciprocal regulation of Akt and Oct4 promotes the self-renewal and survival of embryonal carcinoma cells. Mol. Cell 48, 627–40.

Liu, C., Wang, R., He, Z., Osteil, P., Wilkie, E., Yang, X., Chen, J., Cui, G., Guo, W., Chen, Y., Peng, G., Tam, P.P.L., Jing, N., 2018. Suppressing Nodal Signaling Activity Predisposes Ectodermal Differentiation of Epiblast Stem Cells. Stem cell reports 11, 43–57.

Liu, D., Wang, X., He, D., Sun, C., He, X., Yan, L., Li, Y., Han, J.-D.J., Zheng, P., 2018. Single cell RNA-sequencing reveals the existence of naïve and primed pluripotency in pre-implantation rhesus monkey embryos. Genome Res.

Liu, P., Chen, M., Liu, Y., Qi, L.S., Ding, S., 2018. CRISPR-Based Chromatin Remodeling of the Endogenous Oct4 or Sox2 Locus Enables Reprogramming to Pluripotency. Cell Stem Cell 22, 252–261.e4.

Liu, X., Yao, Y., Ding, H., Han, C., Chen, Y., Zhang, Y., Wang, C., Zhang, X., Zhang, Y., Zhai, Y., Wang, P., Wei, W., Zhang, J., Zhang, L., 2016. USP21 deubiquitylates Nanog to regulate protein stability and stem cell pluripotency. Signal Transduct. Target. Ther. 1, 16024.

Loffreda, A., Jacchetti, E., Antunes, S., Rainone, P., Daniele, T., Morisaki, T., Bianchi, M.E., Tacchetti, C., Mazza, D., 2017. Live-cell p53 single-molecule binding is modulated by C-terminal acetylation and correlates with transcriptional activity. Nat. Commun. 8, 313.

Loh, K.M., Lim, B., 2011. A precarious balance: pluripotency factors as lineage specifiers. Cell Stem Cell 8, 363–9.

Loh, Y.-H., Wu, Q., Chew, J.-L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., Wong, K.-Y., Sung, K.W., Lee, C.W.H., Zhao, X.-D., Chiu, K.-P., Lipovich, L., Kuznetsov, V.A., Robson, P., Stanton, L.W., Wei, C.-L., Ruan, Y., Lim, B., Ng, H.-H., 2006. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat. Genet. 38, 431–40.

López-Martínez, A., Chang, D.T., Chiang, C., Porter, J.A., Ros, M.A., Simandl, B.K., Beachy, P.A., Fallon, J.F., 1995. Limb-patterning activity and restricted posterior localization of the amino-terminal product of Sonic hedgehog cleavage. Curr. Biol. 5, 791–6.

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550.

Luikenhuis, S., Wutz, A., Jaenisch, R., 2001. Antisense transcription through the Xist locus mediates Tsix function in embryonic stem cells. Mol. Cell. Biol. 21, 8512–20.

Macdonald, K.B., Juriloff, D.M., Harris, M.J., 1989. Developmental study of neural tube closure in a mouse stock with a high incidence of exencephaly. Teratology 39, 195–213.

Mantalenakis, S.J., Ketchel, M.M., 1966. Frequency and extent of delayed implantation in lactating rats and mice. J. Reprod. Fertil. 12, 391–4.

Marahrens, Y., Loring, J., Jaenisch, R., 1998. Role of the Xist gene in X chromosome choosing. Cell 92, 657–64.

Marks, H., Kalkan, T., Menafra, R., Denissov, S., Jones, K., Hofemeister, H., Nichols, J., Kranz, A., Stewart, A.F., Smith, A., Stunnenberg, H.G., 2012. The transcriptional and epigenomic foundations of ground state pluripotency. Cell 149, 590–604.

Marson, A., Levine, S.S., Cole, M.F., Frampton, G.M., Brambrink, T., Johnstone, S., Guenther, M.G., Johnston, W.K., Wernig, M., Newman, J., Calabrese, J.M., Dennis, L.M., Volkert, T.L., Gupta, S., Love, J., Hannett, N., Sharp, P.A., Bartel, D.P., Jaenisch, R., Young, R.A., 2008. Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell 134, 521–33.

Martello, G., Bertone, P., Smith, A., 2013. Identification of the missing pluripotency mediator downstream of leukaemia inhibitory factor. EMBO J. 32, 2561–74.

Martello, G., Sugimoto, T., Diamanti, E., Joshi, A., Hannah, R., Ohtsuka, S., Göttgens, B., Niwa, H., Smith, A., 2012. Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. Cell Stem Cell 11, 491–504.

Martin, G.R., 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc. Natl. Acad. Sci. U. S. A. 78, 7634–7638.

Martin, M., 2011. Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads. EMBnet.journal 17, 10–12.

Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A.A., Ko, M.S.H., Niwa, H., 2007. Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. Nat. Cell Biol. 9, 625–35.

Matoba, R., Niwa, H., Masui, S., Ohtsuka, S., Carter, M.G., Sharov, A.A., Ko, M.S.H., 2006. Dissecting Oct3/4-regulated gene networks in embryonic stem cells by expression profiling. PLoS One 1, e26.

Matsuda, T., Nakamura, T., Nakao, K., Arai, T., Katsuki, M., Heike, T., Yokota, T., 1999. STAT3

activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. EMBO J. 18, 4261–4269.

Matsui, Y., 1998. Developmental fates of the mouse germ cell line. Int. J. Dev. Biol. 42, 1037–42.

Matsui, Y., Toksoz, D., Nishikawa, S., Nishikawa, S., Williams, D., Zsebo, K., Hogan, B.L., 1991. Effect of Steel factor and leukaemia inhibitory factor on murine primordial germ cells in culture. Nature 353, 750–2.

Matsui, Y., Zsebo, K., Hogan, B.L., 1992. Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. Cell 70, 841–847.

Mazza, D., Abernathy, A., Golob, N., Morisaki, T., McNally, J.G., 2012. A benchmark for chromatin binding measurements in live cells. Nucleic Acids Res. 40, e119.

Medeiros, R.B., Papenfuss, K.J., Hoium, B., Coley, K., Jadrich, J., Goh, S.-K., Elayaperumal, A., Herrera, J.E., Resnik, E., Ni, H.-T., 2009. Novel sequential ChIP and simplified basic ChIP protocols for promoter co-occupancy and target gene identification in human embryonic stem cells. BMC Biotechnol. 9, 59.

Meilhac, S.M., Adams, R.J., Morris, S.A., Danckaert, A., Le Garrec, J.-F., Zernicka-Goetz, M., 2009. Active cell movements coupled to positional induction are involved in lineage segregation in the mouse blastocyst. Dev. Biol. 331, 210–21.

Menchero, S., Rayon, T., Andreu, M.J., Manzanares, M., 2017. Signaling pathways in mammalian preimplantation development: Linking cellular phenotypes to lineage decisions. Dev. Dyn. 246, 245–261.

Merino, F., Ng, C.K.L., Veerapandian, V., Schöler, H.R., Jauch, R., Cojocaru, V., 2014. Structural basis for the SOX-dependent genomic redistribution of OCT4 in stem cell differentiation. Structure 22, 1274–1286.

Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., Meissner, A., 2008. Dissecting direct reprogramming through integrative genomic analysis. Nature 454, 49–55.

Minkovsky, A., Barakat, T.S., Sellami, N., Chin, M.H., Gunhanlar, N., Gribnau, J., Plath, K., 2013. The pluripotency factor-bound intron 1 of Xist is dispensable for X chromosome inactivation and

reactivation in vitro and in vivo. Cell Rep. 3, 905–18.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., Yamanaka, S., 2003. The homeoprotein nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell 113, 631–642.

Molotkov, A., Mazot, P., Brewer, J.R., Cinalli, R.M., Soriano, P., 2017. Distinct Requirements for FGFR1 and FGFR2 in Primitive Endoderm Development and Exit from Pluripotency. Dev. Cell 41, 511–526.e4.

Molyneaux, K.A., Stallock, J., Schaible, K., Wylie, C., 2001. Time-lapse analysis of living mouse germ cell migration. Dev. Biol. 240, 488–98.

Moretto-Zita, M., Jin, H., Shen, Z., Zhao, T., Briggs, S.P., Xu, Y., 2010. Phosphorylation stabilizes Nanog by promoting its interaction with Pin1. Proc. Natl. Acad. Sci. U. S. A. 107, 13312–7.

Morey, C., Arnaud, D., Avner, P., Clerc, P., 2001. Tsix-mediated repression of Xist accumulation is not sufficient for normal random X inactivation. Hum. Mol. Genet. 10, 1403–11.

Morey, C., Navarro, P., Debrand, E., Avner, P., Rougeulle, C., Clerc, P., 2004. The region 3' to Xist mediates X chromosome counting and H3 Lys-4 dimethylation within the Xist gene. EMBO J. 23, 594–604.

Morikawa, M., Koinuma, D., Mizutani, A., Kawasaki, N., Holmborn, K., Sundqvist, A., Tsutsumi, S., Watabe, T., Aburatani, H., Heldin, C.-H., Miyazono, K., 2016. BMP Sustains Embryonic Stem Cell Self-Renewal through Distinct Functions of Different Krüppel-like Factors. Stem cell reports 6, 64–73.

Morisaki, T., Müller, W.G., Golob, N., Mazza, D., McNally, J.G., 2014. Single-molecule analysis of transcription factor binding at transcription sites in live cells. Nat. Commun. 5, 4456.

Mueller, F., Stasevich, T.J., Mazza, D., McNally, J.G., 2013. Quantifying transcription factor kinetics: at work or at play? Crit. Rev. Biochem. Mol. Biol. 48, 492–514.

Mulas, C., Chia, G., Jones, K.A., Hodgson, A.C., Stirparo, G.G., Nichols, J., 2018. Oct4 regulates the embryonic axis and coordinates exit from pluripotency and germ layer specification in the mouse embryo. Development 145.

Mulas, C., Kalkan, T., Smith, A., 2017. NODAL Secures Pluripotency upon Embryonic Stem Cell

Progression from the Ground State. Stem cell reports 9, 77–91.

Muñoz Descalzo, S., Rué, P., Faunes, F., Hayward, P., Jakt, L.M., Balayo, T., Garcia-Ojalvo, J., Martinez Arias, A., 2013. A competitive protein interaction network buffers Oct4-mediated differentiation to promote pluripotency in embryonic stem cells. Mol. Syst. Biol. 9, 694.

Murayama, H., Masaki, H., Sato, H., Hayama, T., Yamaguchi, T., Nakauchi, H., 2015. Successful reprogramming of epiblast stem cells by blocking nuclear localization of β-catenin. Stem cell reports 4, 103–13.

Nagy, A., 2013. Secondary cell reprogramming systems: as years go by. Curr. Opin. Genet. Dev. 23, 534–9.

Najm, F.J., Chenoweth, J.G., Anderson, P.D., Nadeau, J.H., Redline, R.W., McKay, R.D.G., Tesar, P.J., 2011. Isolation of epiblast stem cells from preimplantation mouse embryos. Cell Stem Cell 8, 318–25.

Nakamura, T., Okamoto, I., Sasaki, K., Yabuta, Y., Iwatani, C., Tsuchiya, H., Seita, Y., Nakamura, S., Yamamoto, T., Saitou, M., 2016. A developmental coordinate of pluripotency among mice, monkeys and humans. Nature 537, 57–62.

Nakatake, Y., Fukui, N., Iwamatsu, Y., Masui, S., Takahashi, K., Yagi, R., Yagi, K., Miyazaki, J.-I., Matoba, R., Ko, M.S.H., Niwa, H., 2006. Klf4 cooperates with Oct3/4 and Sox2 to activate the Lefty1 core promoter in embryonic stem cells. Mol. Cell. Biol. 26, 7772–7782.

Navarro, P., Chambers, I., Karwacki-Neisius, V., Chureau, C., Morey, C., Rougeulle, C., Avner, P., 2008. Molecular coupling of Xist regulation and pluripotency. Science 321, 1693–5.

Nesterova, T.B., Barton, S.C., Surani, M.A., Brockdorff, N., 2001. Loss of Xist imprinting in diploid parthenogenetic preimplantation embryos. Dev. Biol. 235, 343–50.

Nichols, J., Chambers, I., Taga, T., Smith, A., 2001. Physiological rationale for responsiveness of mouse embryonic stem cells to gp130 cytokines. Development 128, 2333–9.

Nichols, J., Jones, K., Phillips, J.M., Newland, S.A., Roode, M., Mansfield, W., Smith, A., Cooke, A., 2009. Validated germline-competent embryonic stem cell lines from nonobese diabetic mice. Nat. Med. 15, 814–818.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Schöler, H.,

Smith, A., 1998. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell 95, 379–391.

Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T., Kanemaki, M., 2009. An auxin-based degron system for the rapid depletion of proteins in nonplant cells. Nat. Methods 6, 917–22.

Nishioka, N., Inoue, K., Adachi, K., Kiyonari, H., Ota, M., Ralston, A., Yabuta, N., Hirahara, S., Stephenson, R.O., Ogonuki, N., Makita, R., Kurihara, H., Morin-Kensicki, E.M., Nojima, H., Rossant, J., Nakao, K., Niwa, H., Sasaki, H., 2009. The Hippo signaling pathway components Lats and Yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass. Dev. Cell 16, 398–410.

Nishioka, N., Yamamoto, S., Kiyonari, H., Sato, H., Sawada, A., Ota, M., Nakao, K., Sasaki, H., 2008. Tead4 is required for specification of trophectoderm in pre-implantation mouse embryos. Mech. Dev. 125, 270–83.

Niwa, H., 2018. The principles that govern transcription factor network functions in stem cells. Development 145.

Niwa, H., Burdon, T., Chambers, I., Smith, A., 1998. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. Genes Dev. 12, 2048–60.

Niwa, H., Masui, S., Chambers, I., Smith, A.G., Miyazaki, J., 2002. Phenotypic complementation establishes requirements for specific POU domain and generic transactivation function of Oct-3/4 in embryonic stem cells. Mol. Cell. Biol. 22, 1526–36.

Niwa, H., Miyazaki, J., Smith, A.G., 2000. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat. Genet. 24, 372–376.

Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., Rossant, J., 2005. Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. Cell 123, 917–29.

Novo, C.L., Javierre, B.-M., Cairns, J., Segonds-Pichon, A., Wingett, S.W., Freire-Pritchett, P., Furlan-Magaril, M., Schoenfelder, S., Fraser, P., Rugg-Gunn, P.J., 2018. Long-Range Enhancer Interactions Are Prevalent in Mouse Embryonic Stem Cells and Are Reorganized upon Pluripotent State Transition. Cell Rep. 22, 2615–2627.

Ogawa, K., Saito, A., Matsui, H., Suzuki, H., Ohtsuka, S., Shimosato, D., Morishita, Y., Watabe, T., Niwa, H., Miyazono, K., 2007. Activin-Nodal signaling is involved in propagation of mouse embryonic stem cells. J. Cell Sci. 120, 55–65.

Ogawa, Y., Lee, J.T., 2003. Xite, X-inactivation intergenic transcription elements that regulate the probability of choice. Mol. Cell 11, 731–43.

Okita, K., Ichisaka, T., Yamanaka, S., 2007. Generation of germline-competent induced pluripotent stem cells. Nature 448, 313–317.

Osorno, R., Tsakiridis, A., Wong, F., Cambray, N., Economou, C., Wilkie, R., Blin, G., Scotting, P.J., Chambers, I., Wilson, V., 2012. The developmental dismantling of pluripotency is reversed by ectopic Oct4 expression. Development 139, 2288–98.

Panning, B., Dausman, J., Jaenisch, R., 1997. X chromosome inactivation is mediated by Xist RNA stabilization. Cell 90, 907–16.

Papaioannou, V.E., McBurney, M.W., Gardner, R.L., Evans, M.J., 1975. Fate of teratocarcinoma cells injected into early mouse embryos. Nature 258, 70–73.

Pardo, M., Lang, B., Yu, L., Prosser, H., Bradley, A., Babu, M.M., Choudhary, J., 2010. An Expanded Oct4 Interaction Network: Implications for Stem Cell Biology, Development, and Disease. Cell Stem Cell 6, 382–395.

Penny, G.D., Kay, G.F., Sheardown, S.A., Rastan, S., Brockdorff, N., 1996. Requirement for Xist in X chromosome inactivation. Nature 379, 131–7.

Pereira, L., Yi, F., Merrill, B.J., 2006. Repression of Nanog gene transcription by Tcf3 limits embryonic stem cell self-renewal. Mol. Cell. Biol. 26, 7479–91.

Pevny, L.H., Sockanathan, S., Placzek, M., Lovell-Badge, R., 1998. A role for SOX1 in neural determination. Development 125, 1967–78.

Plusa, B., Piliszek, A., Frankenberg, S., Artus, J., Hadjantonakis, A.-K., 2008. Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. Development 135, 3081–91.

Pratt, H.P., Ziomek, C.A., Reeve, W.J., Johnson, M.H., 1982. Compaction of the mouse embryo: an analysis of its components. J. Embryol. Exp. Morphol. 70, 113–32.

Quinlan, G.A., Williams, E.A., Tan, S.S., Tam, P.P., 1995. Neuroectodermal fate of epiblast cells in the distal region of the mouse egg cylinder: implication for body plan organization during early embryogenesis. Development 121, 87–98.

Radzisheuskaya, A., Chia, G.L. Bin, dos Santos, R.L., Theunissen, T.W., Castro, L.F.C., Nichols, J., Silva, J.C.R., 2013. A defined Oct4 level governs cell state transitions of pluripotency entry and differentiation into all embryonic lineages. Nat. Cell Biol. 15, 579–90.

Radzisheuskaya, A., Silva, J.C.R.R., 2014. Do all roads lead to Oct4? The emerging concepts of induced pluripotency. Trends Cell Biol. 24, 275–284.

Raggioli, A., Junghans, D., Rudloff, S., Kemler, R., 2014. Beta-catenin is vital for the integrity of mouse embryonic stem cells. PLoS One 9, e86691.

Raj, A., Peskin, C.S., Tranchina, D., Vargas, D.Y., Tyagi, S., 2006. Stochastic mRNA synthesis in mammalian cells. PLoS Biol. 4, e309.

Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., Manke, T., 2016. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 44, W160–W165.

Ramos-Mejía, V., Escalante-Alcalde, D., Kunath, T., Ramírez, L., Gertsenstein, M., Nagy, A., Lomelí, H., 2005. Phenotypic analyses of mouse embryos with ubiquitous expression of Oct4: Effects on mid-hindbrain patterning and gene expression. Dev. Dyn. 232, 180–190.

Rastan, S., Kaufman, M.H., Handyside, A.H., Lyon, M.F., 1980. X-chromosome inactivation in extra-embryonic membranes of diploid parthenogenetic mouse embryos demonstrated by differential staining. Nature 288, 172–3.

Rastan, S., Robertson, E.J., 1985. X-chromosome deletions in embryo-derived (EK) cell lines associated with lack of X-chromosome inactivation. J. Embryol. Exp. Morphol. 90, 379–88.

Reik, W., Walter, J., 2001. Genomic imprinting: parental influence on the genome. Nat. Rev. Genet. 2, 21–32.

Reményi, A., Lins, K., Nissen, L.J., Reinbold, R., Schöler, H.R., Wilmanns, M., 2003. Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers. Genes Dev. 17, 2048–59.

Resnick, J.L., Bixler, L.S., Cheng, L., Donovan, P.J., 1992. Long-term proliferation of mouse primordial germ cells in culture. Nature 359, 550–1.

Riddle, R.D., Johnson, R.L., Laufer, E., Tabin, C., 1993. Sonic hedgehog mediates the polarizing activity of the ZPA. Cell 75, 1401–16.

Rodda, D.J., Chew, J.L., Lim, L.H., Loh, Y.H., Wang, B., Ng, H.H., Robson, P., 2005. Transcriptional regulation of Nanog by OCT4 and SOX2. J. Biol. Chem. 280, 24731–24737.

Rory, S., Brown, G., 2011. DiffBind: differential binding analysis of ChIP-seq peak data. Bioconductor.

Rosenthal, M.D., Wishnow, R.M., Sato, G.H., 1970. In vitro growth and differetiation of clonal populations of multipotential mouse clls derived from a transplantable testicular teratocarcinoma. J. Natl. Cancer Inst. 44, 1001–14.

Rosner, M.H., Vigano, M.A., Ozato, K., Timmons, P.M., Poirier, F., Rigby, P.W., Staudt, L.M., 1990. A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. Nature 345, 686–92.

Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D., Gojis, O., Ellis, I.O., Green, A.R., Ali, S., Chin, S.-F., Palmieri, C., Caldas, C., Carroll, J.S., 2012. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature 481, 389–393.

Ruegger, M., Dewey, E., Gray, W.M., Hobbie, L., Turner, J., Estelle, M., 1998. The TIR1 protein of Arabidopsis functions in auxin response and is related to human SKP2 and yeast grr1p. Genes Dev. 12, 198–207.

Sado, T., Li, E., Sasaki, H., 2002. Effect of TSIX disruption on XIST expression in male ES cells. Cytogenet. Genome Res. 99, 115–8.

Sanchez-Ripoll, Y., Bone, H.K., Owen, T., Guedes, A.M. V, Abranches, E., Kumpfmueller, B., Spriggs, R. V, Henrique, D., Welham, M.J., 2013. Glycogen synthase kinase-3 inhibition enhances translation of pluripotency-associated transcription factors to contribute to maintenance of mouse embryonic stem cell self-renewal. PLoS One 8, e60148.

Saunders, J. W., J. and M.T.G., 1968. Ectodermal-mesodermal interactions in the origin of limb symmetry. Ep. Interact. 78–97.

Saxe, J.P., Tomilin, A., Schöler, H.R., Plath, K., Huang, J., 2009. Post-translational regulation of Oct4 transcriptional activity. PLoS One 4, e4467.

Scaffidi, P., Bianchi, M.E., 2001. Spatially precise DNA bending is an essential activity of the sox2 transcription factor. J. Biol. Chem. 276, 47296–302.

Schneider-Poetsch, T., Ju, J., Eyler, D.E., Dang, Y., Bhat, S., Merrick, W.C., Green, R., Shen, B., Liu, J.O., 2010. Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. Nat. Chem. Biol. 6, 209–217.

Schrode, N., Saiz, N., Di Talia, S., Hadjantonakis, A.-K., 2014. GATA6 levels modulate primitive endoderm cell fate choice and timing in the mouse blastocyst. Dev. Cell 29, 454–67.

Schwarz, B.A., Cetinbas, M., Clement, K., Walsh, R.M., Cheloufi, S., Gu, H., Langkabel, J., Kamiya, A., Schorle, H., Meissner, A., Sadreyev, R.I., Hochedlinger, K., 2018. Prospective Isolation of Poised iPSC Intermediates Reveals Principles of Cellular Reprogramming. Cell Stem Cell 23, 289–305.e5.

Semrau, S., Goldmann, J.E., Soumillon, M., Mikkelsen, T.S., Jaenisch, R., van Oudenaarden, A., 2017. Dynamics of lineage commitment revealed by single-cell transcriptomics of differentiating embryonic stem cells. Nat. Commun. 8, 1096.

Shamblott, M.J., Axelman, J., Wang, S., Bugg, E.M., Littlefield, J.W., Donovan, P.J., Blumenthal, P.D., Huggins, G.R., Gearhart, J.D., 1998. Derivation of pluripotent stem cells from cultured human primordial germ cells. Proc. Natl. Acad. Sci. U. S. A. 95, 13726–31.

Sheardown, S.A., Duthie, S.M., Johnston, C.M., Newall, A.E., Formstone, E.J., Arkell, R.M., Nesterova, T.B., Alghisi, G.C., Rastan, S., Brockdorff, N., 1997. Stabilization of Xist RNA mediates initiation of X chromosome inactivation. Cell 91, 99–107.

Shi, W., Wang, H., Pan, G., Geng, Y., Guo, Y., Pei, D., 2006. Regulation of the pluripotency marker Rex-1 by Nanog and Sox2. J. Biol. Chem. 281, 23319–25.

Shimosato, D., Shiki, M., Niwa, H., 2007. Extra-embryonic endoderm cells derived from ES cells induced by GATA factors acquire the character of XEN cells. BMC Dev. Biol. 7, 80.

Shimozaki, K., Nakashima, K., Niwa, H., Taga, T., 2003. Involvement of Oct3/4 in the enhancement of neuronal differentiation of ES cells in neurogenesis-inducing cultures. Development 130,

2505–12.

Shu, J., Wu, C., Wu, Y., Li, Z., Shao, S., Zhao, W., Tang, X., Yang, H., Shen, L., Zuo, X., Yang, W., Shi, Y., Chi, X., Zhang, H., Gao, G., Shu, Y., Yuan, K., He, W., Tang, C., Zhao, Y., Deng, H., 2013. Induction of pluripotency in mouse somatic cells with lineage specifiers. Cell 153, 963–75.

Silva, J., Barrandon, O., Nichols, J., Kawaguchi, J., Theunissen, T.W., Smith, A., 2008. Promotion of Reprogramming to Ground State Pluripotency by Signal Inhibition. PLoS Biol. 6, e253.

Silva, J., Nichols, J., Theunissen, T.W., Guo, G., van Oosten, A.L., Barrandon, O., Wray, J., Yamanaka, S., Chambers, I., Smith, A., 2009. Nanog Is the Gateway to the Pluripotent Ground State. Cell 138, 722–37.

Sim, Y.-J., Kim, M.-S., Nayfeh, A., Yun, Y.-J., Kim, S.-J., Park, K.-T., Kim, C.-H., Kim, K.-S., 2017. 2i Maintains a Naive Ground State in ESCs through Two Distinct Epigenetic Mechanisms. Stem cell reports 8, 1312–1328.

Simandi, Z., Horvath, A., Wright, L.C., Cuaranta-Monroy, I., De Luca, I., Karolyi, K., Sauer, S., Deleuze, J.-F., Gudas, L.J., Cowley, S.M., Nagy, L., 2016. OCT4 Acts as an Integrator of Pluripotency and Signal-Induced Differentiation. Mol. Cell 63, 647–661.

Simmler, M.C., Cattanach, B.M., Rasberry, C., Rougeulle, C., Avner, P., 1993. Mapping the murine Xce locus with (CA)n repeats. Mamm. Genome 4, 523–30.

Singh, A.M., Hamazaki, T., Hankowski, K.E., Terada, N., 2007. A heterogeneous expression pattern for Nanog in embryonic stem cells. Stem Cells 25, 2534–42.

Skinner, S.O., Xu, H., Nagarkar-Jaiswal, S., Freire, P.R., Zwaka, T.P., Golding, I., 2016. Single-cell analysis of transcription kinetics across the cell cycle. Elife 5, e12175.

Skowyra, D., Craig, K.L., Tyers, M., Elledge, S.J., Harper, J.W., 1997. F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex. Cell 91, 209–19.

Smith, A., 2017. Formative pluripotency: the executive phase in a developmental continuum. Development 144, 365–373.

Smith, A.G., Heath, J.K., Donaldson, D.D., Wong, G.G., Moreau, J., Stahl, M., Rogers, D., 1988. Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. Nature 336, 688–90.

Smith, Z.D., Chan, M.M., Mikkelsen, T.S., Gu, H., Gnirke, A., Regev, A., Meissner, A., 2012. A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature 484, 339–44.

Song, L., Chen, J., Peng, G., Tang, K., Jing, N., 2016. Dynamic Heterogeneity of Brachyury in Mouse Epiblast Stem Cells Mediates Distinct Response to Extrinsic Bone Morphogenetic Protein (BMP) Signaling. J. Biol. Chem. 291, 15212–25.

Soufi, A., Garcia, M.F., Jaroszewicz, A., Osman, N., Pellegrini, M., Zaret, K.S., 2015. Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. Cell 161, 555–568.

Sousa, E.J., Stuart, H.T., Bates, L.E., Ghorbani, M., Nichols, J., Dietmann, S., Silva, J.C.R., 2018. Exit from Naive Pluripotency Induces a Transient X Chromosome Inactivation-like State in Males. Cell Stem Cell 22, 919–928.e6.

Sozen, B., Amadei, G., Cox, A., Wang, R., Na, E., Czukiewska, S., Chappell, L., Voet, T., Michel, G., Jing, N., Glover, D.M., Zernicka-Goetz, M., 2018. Self-assembly of embryonic and two extra-embryonic stem cell types into gastrulating embryo-like structures. Nat. Cell Biol. 20, 979–989.

Starmer, J., Magnuson, T., 2009. A new model for random X chromosome inactivation. Development 136, 1–10.

Stavropoulos, N., Lu, N., Lee, J.T., 2001. A functional role for Tsix transcription in blocking Xist RNA accumulation but not in X-chromosome choice. Proc. Natl. Acad. Sci. U. S. A. 98, 10232–7.

Stevens, L.C., 1970. The development of transplantable teratocarcinomas from intratesticular grafts of pre- and postimplantation mouse embryos. Dev. Biol. 21, 364–82.

Stevens, L.C., 1967. Origin of testicular teratomas from primordial germ cells in mice. J. Natl. Cancer Inst. 38, 549–52.

Stevens, L.C., Hummel, K.P., 1957. A description of spontaneous congenital testicular teratomas in strain 129 mice. J. Natl. Cancer Inst. 18, 719–47.

Stevens, L.C., Little, C.C., 1954. Spontaneous Testicular Teratomas in an Inbred Strain of Mice. Proc. Natl. Acad. Sci. U. S. A. 40, 1080–7.

Stewart, C.L., Kaspar, P., Brunet, L.J., Bhatt, H., Gadi, I., Köntgen, F., Abbondanzo, S.J., 1992. Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. Nature

359, 76–9.

Stewart, T.A., Mintz, B., 1981. Successive generations of mice produced from an established culture line of euploid teratocarcinoma cells. Proc. Natl. Acad. Sci. U. S. A. 78, 6314–8.

Stirparo, G.G., Boroviak, T., Guo, G., Nichols, J., Smith, A., Bertone, P., 2018. Integrated analysis of single-cell embryo data yields a unified transcriptome signature for the human pre-implantation epiblast. Development 145.

Strebinger, D., Friman, E.T., Deluz, C., Govindan, S., Alber, A.B., Suter, D.M., 2018. Endogenous fluctuations of OCT4 and SOX2 bias pluripotent cell fate decisions. bioRxiv.

Stresemann, C., Lyko, F., 2008. Modes of action of the DNA methyltransferase inhibitors azacytidine and decitabine. Int. J. cancer 123, 8–13.

Stuart, H.T., Van Oosten, A.L., Radzisheuskaya, A., Martello, G., Miller, A., Dietmann, S., Nichols, J., Silva, J.C.R., 2014. NANOG amplifies STAT3 activation and they synergistically induce the naive pluripotent program. Curr. Biol. 24, 340–346.

Summerbell, D., 1983. The effect of local application of retinoic acid to the anterior margin of the developing chick limb. J. Embryol. Exp. Morphol. 78, 269–89.

Surani, M.A., Hayashi, K., Hajkova, P., 2007. Genetic and epigenetic regulators of pluripotency. Cell 128, 747–62.

Suter, D.M., Molina, N., Gatfield, D., Schneider, K., Schibler, U., Naef, F., 2011. Mammalian genes are transcribed with widely different bursting kinetics. Science 332, 472–4.

Suter, D.M., Tirefort, D., Julien, S., Krause, K.-H., 2009. A Sox1 to Pax6 switch drives neuroectoderm to radial glia progression during differentiation of mouse embryonic stem cells. Stem Cells 27, 49–58.

Tada, T., Takagi, N., 1992. Early development and X-chromosome inactivation in mouse parthenogenetic embryos. Mol. Reprod. Dev. 31, 20–7.

Takahashi, K., Yamanaka, S., 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126, 663–76.

Takaoka, K., Yamamoto, M., Hamada, H., 2011. Origin and role of distal visceral endoderm, a group

of cells that determines anterior-posterior polarity of the mouse embryo. Nat. Cell Biol. 13, 743–52.

Takaoka, K., Yamamoto, M., Shiratori, H., Meno, C., Rossant, J., Saijoh, Y., Hamada, H., 2006. The mouse embryo autonomously acquires anterior-posterior polarity at implantation. Dev. Cell 10, 451–9.

Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficz, G., Krueger, F., Oxley, D., Santos, F., Clarke, J., Mansfield, W., Reik, W., Bertone, P., Smith, A., 2014. Resetting transcription factor control circuitry toward ground-state pluripotency in human. Cell 158, 1254–69.

Tam, P.P., Beddington, R.S., 1987. The formation of mesodermal tissues in the mouse embryo during gastrulation and early organogenesis. Development 99, 109–26.

Tam, P.P., Behringer, R.R., 1997. Mouse gastrulation: the formation of a mammalian body plan. Mech. Dev. 68, 3–25.

Tam, P.P., Parameswaran, M., Kinder, S.J., Weinberger, R.P., 1997. The allocation of epiblast cells to the embryonic heart and other mesodermal lineages: the role of ingression and tissue movement during gastrulation. Development 124, 1631–42.

Tam, P.P., Snow, M.H., 1981. Proliferation and migration of primordial germ cells during compensatory growth in mouse embryos. J. Embryol. Exp. Morphol. 64, 133–47.

Tam, P.P.L., Loebel, D. a F., 2007. Gene function in mouse embryogenesis: get set for gastrulation. Nat. Rev. Genet. 8, 368–81.

Tan, X., Calderon-Villalobos, L.I.A., Sharon, M., Zheng, C., Robinson, C. V, Estelle, M., Zheng, N., 2007. Mechanism of auxin perception by the TIR1 ubiquitin ligase. Nature 446, 640–645.

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., McKay, R.D.G., 2007. New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature 448, 196–9.

Theunissen, T.W., Costa, Y., Radzisheuskaya, A., van Oosten, A.L., Lavial, F., Pain, B., Castro, L.F.C., Silva, J.C.R., 2011. Reprogramming capacity of Nanog is functionally conserved in vertebrates and resides in a unique homeodomain. Development 138, 4853–4865.

Theunissen, T.W., Friedli, M., He, Y., Planet, E., O'Neil, R.C., Markoulaki, S., Pontis, J., Wang, H.,

Iouranova, A., Imbeault, M.M., Duc, J., Cohen, M.A., Wert, K.J., Castanon, R., Zhang, Z., Huang, Y., Nery, J.R., Drotar, J., Lungjangwa, T., Trono, D., Ecker, J.R., Jaenisch, R., 2016. Molecular Criteria for Defining the Naive Human Pluripotent State. Cell Stem Cell 19, 502–515.

Theunissen, T.W., Powell, B.E., Wang, H., Mitalipova, M., Faddah, D.A., Reddy, J., Fan, Z.P., Maetzel, D., Ganz, K., Shi, L., Lungjangwa, T., Imsoonthornruksa, S., Stelzer, Y., Rangarajan, S., D'Alessio, A., Zhang, J., Gao, Q., Dawlaty, M.M., Young, R.A., Gray, N.S., Jaenisch, R., 2014. Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. Cell Stem Cell 15, 471–87.

Theunissen, T.W., Van Oosten, A.L., Castelo-Branco, G., Hall, J., Smith, A., Silva, J.C.R., 2011. Nanog overcomes reprogramming barriers and induces pluripotency in minimal conditions. Curr. Biol. 21, 65–71.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., Jones, J.M., 1998. Embryonic stem cell lines derived from human blastocysts. Science 282, 1145–7.

Thomson, J.A., Kalishman, J., Golos, T.G., Durning, M., Harris, C.P., Becker, R.A., Hearn, J.P., 1995. Isolation of a primate embryonic stem cell line. Proc. Natl. Acad. Sci. U. S. A. 92, 7844–8.

Thomson, J.A., Kalishman, J., Golos, T.G., Durning, M., Harris, C.P., Hearn, J.P., 1996. Pluripotent cell lines derived from common marmoset (Callithrix jacchus) blastocysts. Biol. Reprod. 55, 254–9.

Thomson, M., Liu, S.J., Zou, L.-N., Smith, Z., Meissner, A., Ramanathan, S., 2011. Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. Cell 145, 875–89.

Tickle, C., Alberts, B., Wolpert, L., Lee, J., 1982. Local application of retinoic acid to the limb bond mimics the action of the polarizing region. Nature 296, 564–6.

Tickle, C., Summerbell, D., Wolpert, L., 1975. Positional signalling and specification of digits in chick limb morphogenesis. Nature 254, 199–202.

Tokuzawa, Y., Kaiho, E., Maruyama, M., Takahashi, K., Mitsui, K., Maeda, M., Niwa, H., Yamanaka, S., 2003. Fbx15 is a novel target of Oct3/4 but is dispensable for embryonic stem cell self-renewal and mouse development. Mol. Cell. Biol. 23, 2699–708.

Torres-Padilla, M.-E., Chambers, I., 2014. Transcription factor heterogeneity in pluripotent stem cells:

a stochastic advantage. Development 141, 2173–81.

Toyooka, Y., Shimosato, D., Murakami, K., Takahashi, K., Niwa, H., 2008. Identification and characterization of subpopulations in undifferentiated ES cell culture. Development 135, 909–18.

Trott, J., Martinez Arias, A., 2013. Single cell lineage analysis of mouse embryonic stem cells at the exit from pluripotency. Biol. Open 2, 1049–56.

Tsai, S.-Y., Bouwman, B.A., Ang, Y.-S., Kim, S.J., Lee, D.-F., Lemischka, I.R., Rendl, M., 2011. Single transcription factor reprogramming of hair follicle dermal papilla cells to induced pluripotent stem cells. Stem Cells 29, 964–71.

Tsakiridis, A., Huang, Y., Blin, G., Skylaki, S., Wymeersch, F., Osorno, R., Economou, C., Karagianni, E., Zhao, S., Lowell, S., Wilson, V., 2014. Distinct Wnt-driven primitive streak-like populations reflect in vivo lineage precursors. Development 141, 1209–21.

Tsuji-Takayama, K., Inoue, T., Ijiri, Y., Otani, T., Motoda, R., Nakamura, S., Orita, K., 2004. Demethylating agent, 5-azacytidine, reverses differentiation of embryonic stem cells. Biochem. Biophys. Res. Commun. 323, 86–90.

Turnpenny, L., Brickwood, S., Spalluto, C.M., Piper, K., Cameron, I.T., Wilson, D.I., Hanley, N.A., 2003. Derivation of human embryonic germ cells: an alternative source of pluripotent stem cells. Stem Cells 21, 598–609.

Vallier, L., Alexander, M., Pedersen, R.A., 2005. Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. J. Cell Sci. 118, 4495–509.

van den Berg, D.L.C., Zhang, W., Yates, A., Engelen, E., Takacs, K., Bezstarosti, K., Demmers, J., Chambers, I., Poot, R.A., 2008. Estrogen-related receptor beta interacts with Oct4 to positively regulate Nanog gene expression. Mol. Cell. Biol. 28, 5986–95.

van den Brink, S.C., Baillie-Johnson, P., Balayo, T., Hadjantonakis, A.-K., Nowotschin, S., Turner, D.A., Martinez Arias, A., 2014. Symmetry breaking, germ layer specification and axial organisation in aggregates of mouse embryonic stem cells. Development 141, 4231–42.

Velkey, J.M., O'Shea, K.S., 2003. Oct4 RNA interference induces trophectoderm differentiation in mouse embryonic stem cells. Genesis 37, 18–24.

Vestweber, D., Gossler, A., Boller, K., Kemler, R., 1987. Expression and distribution of cell adhesion molecule uvomorulin in mouse preimplantation embryos. Dev. Biol. 124, 451–6.

Vigneau, S., Augui, S., Navarro, P., Avner, P., Clerc, P., 2006. An essential role for the DXPas34 tandem repeat and Tsix transcription in the counting process of X chromosome inactivation. Proc. Natl. Acad. Sci. U. S. A. 103, 7390–5.

Vinot, S., Le, T., Ohno, S., Pawson, T., Maro, B., Louvet-Vallée, S., 2005. Asymmetric distribution of PAR proteins in the mouse embryo begins at the 8-cell stage during compaction. Dev. Biol. 282, 307–19.

Waddington, C.H., 1957. The strategy of the genes: a discussion of some aspects of theoretical biology. Allen & Unwin.

Wallace, M.E., Knights, P.J., Anderson, J.R., 1978. Inheritance and morphology of exencephaly, a neonatal lethal recessive with partial penetrance, in the house mouse. Genet. Res. 32, 135–49.

Wang, G., Zhang, H., Zhao, Y., Li, J., Cai, J., Wang, P., Meng, S., Feng, J., Miao, C., Ding, M., Li, D., Deng, H., 2005. Noggin and bFGF cooperate to maintain the pluripotency of human embryonic stem cells in the absence of feeder layers. Biochem. Biophys. Res. Commun. 330, 934–42.

Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D.N., Theunissen, T.W., Orkin, S.H., 2006. A protein interaction network for pluripotency of embryonic stem cells. Nature 444, 364–8.

Wang, T., Chen, K., Zeng, X., Yang, J., Wu, Y., Shi, X., Qin, B., Zeng, L., Esteban, M.A., Pan, G., Pei, D., 2011. The histone demethylases Jhdm1a/1b enhance somatic cell reprogramming in a vitamin-C-dependent manner. Cell Stem Cell 9, 575–87.

Ware, C.B., Horowitz, M.C., Renshaw, B.R., Hunt, J.S., Liggitt, D., Koblar, S.A., Gliniak, B.C., McKenna, H.J., Papayannopoulou, T., Thoma, B., 1995. Targeted disruption of the low-affinity leukemia inhibitory factor receptor gene causes placental, skeletal, neural and metabolic defects and results in perinatal death. Development 121, 1283–99.

Ware, C.B., Nelson, A.M., Mecham, B., Hesson, J., Zhou, W., Jonlin, E.C., Jimenez-Caliani, A.J., Deng, X., Cavanaugh, C., Cook, S., Tesar, P.J., Okada, J., Margaretha, L., Sperber, H., Choi, M., Blau, C.A., Treuting, P.M., Hawkins, R.D., Cirulli, V., Ruohola-Baker, H., 2014. Derivation of naive human embryonic stem cells. Proc. Natl. Acad. Sci. U. S. A. 111, 4484–9.

Webb, S., de Vries, T.J., Kaufman, M.H., 1992. The differential staining pattern of the X chromosome in the embryonic and extraembryonic tissues of postimplantation homozygous tetraploid mouse embryos. Genet. Res. 59, 205–14.

Wei, F., Schöler, H.R., Atchison, M.L., 2007. Sumoylation of Oct4 enhances its stability, DNA binding, and transactivation. J. Biol. Chem. 282, 21551–21560.

Weitlauf, H.M., Greenwald, G.S., 1968. Survival of blastocysts in the uteri of ovariectomized mice. J. Reprod. Fertil. 17, 515–20.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., Young, R.A., 2013. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell 153, 307–319.

Wicklow, E., Blij, S., Frum, T., Hirate, Y., Lang, R.A., Sasaki, H., Ralston, A., 2014. HIPPO pathway members restrict SOX2 to the inner cell mass where it promotes ICM fates in the mouse blastocyst. PLoS Genet. 10, e1004618.

Williams, M., Burdsal, C., Periasamy, A., Lewandoski, M., Sutherland, A., 2012. Mouse primitive streak forms in situ by initiation of epithelial to mesenchymal transition without migration of a cell population. Dev. Dyn. 241, 270–83.

Williams, R.L., Hilton, D.J., Pease, S., Willson, T.A., Stewart, C.L., Gearing, D.P., Wagner, E.F., Metcalf, D., Nicola, N.A., Gough, N.M., 1988. Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. Nature 336, 684–687.

Wray, J., Kalkan, T., Gomez-Lopez, S., Eckardt, D., Cook, A., Kemler, R., Smith, A., 2011. Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. Nat. Cell Biol. 13, 838–45.

Wray, J., Kalkan, T., Smith, A.G., 2010. The ground state of pluripotency. Biochem. Soc. Trans. 38, 1027–1032.

Xu, C., Rosler, E., Jiang, J., Lebkowski, J.S., Gold, J.D., O'Sullivan, C., Delavan-Boorsma, K., Mok, M., Bronstein, A., Carpenter, M.K., 2005. Basic fibroblast growth factor supports undifferentiated human embryonic stem cell growth without conditioned medium. Stem Cells 23, 315–23.

Xu, N., Tsai, C.-L., Lee, J.T., 2006. Transient homologous chromosome pairing marks the onset of X

inactivation. Science 311, 1149–52.

Xu, R.-H., Chen, X., Li, D.S., Li, R., Addicks, G.C., Glennon, C., Zwaka, T.P., Thomson, J.A., 2002. BMP4 initiates human embryonic stem cell differentiation to trophoblast. Nat. Biotechnol. 20, 1261–4.

Xu, R.-H., Peck, R.M., Li, D.S., Feng, X., Ludwig, T., Thomson, J.A., 2005. Basic FGF and suppression of BMP signaling sustain undifferentiated proliferation of human ES cells. Nat. Methods 2, 185–90.

Yagi, R., Kohn, M.J., Karavanova, I., Kaneko, K.J., Vullhorst, D., DePamphilis, M.L., Buonanno, A., 2007. Transcription factor TEAD4 specifies the trophectoderm lineage at the beginning of mammalian development. Development 134, 3827–36.

Yamaguchi, S., Kimura, H., Tada, M., Nakatsuji, N., Tada, T., 2005. Nanog expression in mouse germ cell development. Gene Expr. Patterns 5, 639–46.

Yamanaka, Y., Lanner, F., Rossant, J., 2010. FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. Development 137, 715–24.

Yang, J., Ryan, D.J., Wang, W., Tsang, J.C.-H., Lan, G., Masaki, H., Gao, X., Antunes, L., Yu, Y., Zhu, Z., Wang, J., Kolodziejczyk, A.A., Campos, L.S., Wang, C., Yang, F., Zhong, Z., Fu, B., Eckersley-Maslin, M.A., Woods, M., Tanaka, Y., Chen, X., Wilkinson, A.C., Bussell, J., White, J., Ramirez-Solis, R., Reik, W., Göttgens, B., Teichmann, S.A., Tam, P.P.L., Nakauchi, H., Zou, X., Lu, L., Liu, P., 2017. Establishment of mouse expanded potential stem cells. Nature 550, 393–397.

Yang, J., van Oosten, A.L., Theunissen, T.W., Guo, G., Silva, J.C.R., Smith, A., 2010. Stat3 activation is limiting for reprogramming to ground state pluripotency. Cell Stem Cell 7, 319–28.

Yang, S.-H., Kalkan, T., Morissroe, C., Marks, H., Stunnenberg, H., Smith, A., Sharrocks, A.D., 2014. Otx2 and Oct4 drive early enhancer activation during embryonic stem cell transition from naive pluripotency. Cell Rep. 7, 1968–81.

Yao, Y., Lu, Y., Chen, W.-C., Jiang, Y., Cheng, T., Ma, Y., Lu, L., Dai, W., 2014. Cobalt and nickel stabilize stem cell transcription factor OCT4 through modulating its sumoylation and ubiquitination. PLoS One 9, e86620.

Ye, S., Li, P., Tong, C., Ying, Q.-L., 2013. Embryonic stem cell self-renewal pathways converge on the

transcription factor Tfcp2l1. EMBO J. 32, 2548–60.

Yeom, Y.I., Fuhrmann, G., Ovitt, C.E., Brehm, A., Ohbo, K., Gross, M., Hübner, K., Schöler, H.R., 1996. Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. Development 122, 881–94.

Yesudhas, D., Anwar, M.A., Panneerselvam, S., Durai, P., Shah, M., Choi, S., 2016. Structural Mechanism behind Distinct Efficiency of Oct4/Sox2 Proteins in Differentially Spaced DNA Complexes. PLoS One 11, e0147240.

Yi, F., Pereira, L., Merrill, B.J., 2008. Tcf3 functions as a steady-state limiter of transcriptional programs of mouse embryonic stem cell self-renewal. Stem Cells 26, 1951–60.

Yin, R., Mao, S.-Q., Zhao, B., Chong, Z., Yang, Y., Zhao, C., Zhang, D., Huang, H., Gao, J., Li, Z., Jiao, Y., Li, C., Liu, S., Wu, D., Gu, W., Yang, Y.-G., Xu, G.-L., Wang, H., 2013. Ascorbic acid enhances Tet-mediated 5-methylcytosine oxidation and promotes DNA demethylation in mammals. J. Am. Chem. Soc. 135, 10396–403.

Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., Smith, A., 2008. The ground state of embryonic stem cell self-renewal. Nature 453, 519–23.

Ying, Q.L., Nichols, J., Chambers, I., Smith, A., 2003. BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell 115, 281–92.

Yoshida, K., Taga, T., Saito, M., Suematsu, S., Kumanogoh, A., Tanaka, T., Fujiwara, H., Hirata, M., Yamagami, T., Nakahata, T., Hirabayashi, T., Yoneda, Y., Tanaka, K., Wang, W.Z., Mori, C., Shiota, K., Yoshida, N., Kishimoto, T., 1996. Targeted disruption of gp130, a common signal transducer for the interleukin 6 family of cytokines, leads to myocardial and hematological disorders. Proc. Natl. Acad. Sci. U. S. A. 93, 407–11.

Young, R.A., 2011. Control of the embryonic stem cell state. Cell 144, 940–54.

Yuan, H., Corbi, N., Basilico, C., Dailey, L., 1995. Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. Genes Dev. 9, 2635–2645.

Yuan, X., Wan, H., Zhao, X., Zhu, S., Zhou, Q., Ding, S., 2011. Brief report: combined chemical treatment enables Oct4-induced reprogramming from mouse embryonic fibroblasts. Stem Cells

29, 549–53.

Zeineddine, D., Papadimou, E., Chebli, K., Gineste, M., Liu, J., Grey, C., Thurig, S., Behfar, A., Wallace, V.A., Skerjanc, I.S., Pucéat, M., 2006. Oct-3/4 dose dependently regulates specification of embryonic stem cells toward a cardiac lineage and early heart development. Dev. Cell 11, 535–46.

Zhang, H., Gayen, S., Xiong, J., Zhou, B., Shanmugam, A.K., Sun, Y., Karatas, H., Liu, L., Rao, R.C., Wang, S., Nesvizhskii, A.I., Kalantry, S., Dou, Y., 2016. MLL1 Inhibition Reprograms Epiblast Stem Cells to Naive Pluripotency. Cell Stem Cell 18, 481–94.

Zunder, E.R., Lujan, E., Goltsev, Y., Wernig, M., Nolan, G.P., 2015. A continuous molecular roadmap to iPSC reprogramming through progression analysis of single-cell mass cytometry. Cell Stem Cell 16, 323–37.

Zwaka, T.P., 2012. Pluripotency network in embryonic stem cells: maybe Leibniz was right all along. Cell Stem Cell 11, 441–2.

# Appendix A

Unbiased list of expression changes in iCdx2 ESCs from Day 0 to Day 4 of induction. Genes with an adjusted P-value < 0.01 shown, top 200 ordered by log fold change shown.

| Gene | Base mean | log2(FoldChange from D0 to D4) | Adjusted P-value |
|---|---|---|---|
| Iigp1 | 86.62 | 10.49 | 2.62E-17 |
| Vtcn1 | 111.64 | 10.07 | 2.55E-16 |
| Serpinb9g | 122.46 | 9.90 | 7.62E-15 |
| Oasl2 | 68.67 | 9.75 | 6.94E-15 |
| 1810065E05Rik | 39.41 | 9.58 | 9.94E-14 |
| Olr1 | 343.67 | 9.57 | 1.07E-18 |
| C1ql2 | 60.21 | 9.36 | 3.84E-11 |
| Dppa1 | 739.76 | 9.24 | 1.67E-68 |
| 1700071M16Rik | 31.70 | 9.24 | 1.67E-11 |
| Timd2 | 149.02 | 9.22 | 8.70E-14 |
| Serpinb9c | 100.65 | 9.05 | 5.89E-13 |
| Sqor | 57.06 | 9.05 | 8.00E-11 |
| Reg3b | 51.58 | 9.02 | 9.52E-13 |
| Pparg | 337.77 | 9.01 | 2.21E-38 |
| Wfdc17 | 17.09 | 8.99 | 1.24E-10 |
| Acta2 | 152.19 | 8.94 | 2.68128E-22 |
| 1600025M17Rik | 59.62 | 8.90 | 8.37E-11 |
| BC051665 | 1518.17 | 8.86 | 2.49E-88 |
| Postn | 44.45 | 8.82 | 8.63173E-12 |
| Tmprss4 | 29.60 | 8.81 | 8.42E-12 |
| Slc16a2 | 75.04 | 8.81 | 2.75E-12 |
| Pla2g7 | 1376.90 | 8.80 | 9.31E-126 |
| Fthl17a | 137.97 | 8.75 | 3.66E-16 |
| Gucy2c | 77.86 | 8.75 | 8.93E-16 |
| H19 | 47988.60 | 8.70 | 3.99E-63 |
| Tmem54 | 60.96 | 8.65 | 1.13E-11 |
| Gm4926 | 269.91 | 8.59 | 4.33E-29 |
| 1700086L19Rik | 25.33 | 8.57 | 9.94E-11 |
| Gata3 | 352.57 | 8.56 | 9.74E-51 |
| Tlx2 | 246.65 | -8.54 | 2.23E-13 |
| Myo7b | 159.37 | 8.53 | 1.06E-28 |
| Cartpt | 18.24 | 8.50 | 1.57E-08 |
| Noxa1 | 17.80 | 8.48 | 1.29E-10 |
| Sod3 | 17.03 | 8.47 | 5.88E-10 |
| Ctsll3 | 128.51 | 8.46 | 9.25E-11 |
| Scnn1g | 34.67 | 8.45 | 2.34E-09 |
| Igf2 | 6683.27 | 8.43 | 1.86E-41 |
| Add2 | 412.09 | 8.41 | 5.62553E-56 |
| Fnd3c2 | 92.94 | 8.40 | 2.90E-10 |
| Mid2 | 83.75 | 8.38 | 9.87E-14 |
| Fgd3 | 340.64 | 8.38 | 3.91E-39 |
| Mcub | 16.80 | 8.33 | 3.30E-10 |
| Ifit1 | 28.84 | 8.31 | 7.11E-09 |
| BC053393 | 698.18 | 8.28 | 9.54E-96 |
| Flt3 | 40.89 | 8.23 | 2.47E-08 |
| Pde6h | 14.67 | 8.18 | 1.84E-08 |
| Entpd1 | 1317.85 | 8.15 | 2.73E-101 |
| Enpp1 | 706.96 | 8.13 | 4.13E-106 |
| Hand1 | 931.51 | 8.13 | 3.16E-136 |
| Ccbe1 | 41.01 | -8.11 | 3.00E-10 |
| Wdr72 | 59.99 | 8.10 | 9.73E-11 |
| Cdh5 | 1553.08 | 8.07 | 1.19E-92 |
| Duoxa2 | 52.15 | 8.06 | 2.11E-10 |
| Rhox12 | 39.77 | 8.02 | 7.51E-10 |
| Tmprss2 | 1813.23 | 7.97 | 4.24E-23 |
| Epn3 | 554.62 | 7.97 | 4.42E-47 |
| Tmem171 | 72.82 | 7.95 | 1.89E-23 |
| Exoc3l4 | 626.68 | 7.92 | 4.32E-74 |
| Clic5 | 745.68 | 7.91 | 8.05E-41 |
| Enpp2 | 75.82 | 7.88 | 3.45E-17 |
| Mettl7b | 14.61 | 7.84 | 1.94E-08 |
| Tifa | 203.99 | 7.83 | 9.07E-44 |
| Ccrl2 | 30.03 | 7.82 | 1.50E-09 |
| Osm | 156.71 | -7.82 | 1.86E-27 |
| Mgam | 114.36 | 7.82 | 6.46E-39 |
| Pcdh12 | 570.34 | 7.81 | 8.70E-78 |
| Trpv5 | 226.27 | 7.80 | 5.16E-40 |
| Mlkl | 14.49 | 7.80 | 2.46E-07 |
| Gad2 | 17.64 | -7.80 | 2.17E-06 |
| Slc1a2 | 9.30 | 7.79 | 6.00E-08 |
| Pnliprp2 | 11.51 | 7.79 | 4.31E-07 |
| Tmigd1 | 10.22 | 7.77 | 1.36E-07 |
| Dmkn | 397.10 | 7.76 | 1.89E-50 |
| Pkp1 | 40.81 | 7.76 | 1.03E-11 |
| Lgals3bp | 623.28 | 7.75 | 1.43E-81 |
| Serpinb9d | 22.75 | 7.73 | 1.40E-08 |
| Atp8b1 | 257.29 | 7.72 | 2.88E-59 |
| Tgm7 | 7.58 | 7.71 | 9.33E-06 |
| Serpinb9f | 90.65 | 7.70 | 1.46E-12 |
| Slc6a14 | 435.85 | 7.67 | 1.05E-56 |
| Tspan8 | 544.83 | 7.67 | 6.13E-87 |
| Tmem132e | 182.70 | 7.65 | 2.25E-10 |
| Krt18 | 14620.81 | 7.64 | 0.00E+00 |
| Cd200r2 | 7.88 | 7.64 | 5.43E-05 |
| Akr1c19 | 21.27 | 7.63 | 2.90E-07 |
| Wnt5a | 172.38 | 7.62 | 2.93E-58 |
| Ptprb | 106.42 | 7.61 | 2.17E-23 |
| Psg29 | 83.41 | 7.61 | 1.77E-08 |
| Jakmip2 | 7.23 | 7.56 | 1.37E-05 |
| Plac1 | 513.80 | 7.52 | 2.28943E-17 |
| Igf2os | 501.33 | 7.52 | 3.55E-99 |
| Havcr1 | 171.92 | 7.52 | 3.72E-22 |
| Phox2b | 25.60 | 7.51 | 6.44E-11 |
| Slc9a3 | 12.52 | 7.50 | 1.08E-07 |
| P4ha3 | 32.18 | 7.48 | 1.20E-08 |
| Metrnl | 557.08 | 7.48 | 4.84E-74 |
| Prl2c5 | 148.85 | 7.45 | 7.47E-09 |
| Alpk1 | 100.21 | 7.45 | 3.12E-26 |
| Slc16a7 | 55.56 | 7.44 | 3.51E-20 |
| Kcna6 | 8.85 | 7.43 | 4.38E-06 |
| Serpinb9e | 287.66 | 7.43 | 3.96E-42 |
| Gpr17 | 21.53 | 7.41 | 2.19E-08 |
| Slc6a2 | 954.30 | 7.41 | 1.39E-28 |
| Hephl1 | 12.53 | 7.41 | 3.75E-07 |
| Hemgn | 13.19 | 7.41 | 3.93E-07 |
| Tmem252 | 23.80 | 7.40 | 4.97E-10 |
| Lin7a | 6.43 | 7.40 | 1.53E-05 |
| Sp5 | 328.03 | -7.39 | 4.21E-08 |
| Chdh | 324.99 | 7.39 | 1.19E-51 |
| Pla2g4d | 49.99 | 7.38 | 7.20E-11 |
| Mir322 | 34.63 | 7.38 | 7.17E-09 |
| 4930442L01Rik | 14.55 | 7.36 | 1.85E-07 |
| Pou4f3 | 23.30 | 7.36 | 2.68E-10 |
| Chst1 | 271.64 | 7.35 | 5.77E-15 |
| Vgll1 | 75.66 | 7.32 | 2.96E-22 |
| Dhrs9 | 53.14 | 7.31 | 2.07E-10 |
| Gm11190 | 9.26 | -7.31 | 8.70E-06 |
| Mfsd6l | 18.82 | 7.31 | 6.05E-07 |
| Hcls1 | 84.50 | 7.30 | 2.57E-15 |
| Thbs1 | 433.16 | 7.29 | 4.45E-57 |
| Pcdhac2 | 26.07 | 7.27 | 2.71E-10 |
| Gbp2b | 35.28 | 7.25 | 2.85E-08 |
| Prl7d1 | 119.37 | 7.24 | 1.11E-07 |
| Serpinb1b | 10.40 | -7.23 | 1.08E-04 |
| Hoxb5 | 7.54 | 7.21 | 2.00E-05 |
| Nxf7 | 4846.84 | 7.21 | 1.69E-30 |
| Isg15 | 12.64 | 7.21 | 6.94E-06 |
| Itga2 | 58.62 | 7.20 | 1.05E-07 |
| Irx4 | 91.35 | 7.18 | 3.10E-32 |
| Nav3 | 171.78 | 7.17 | 1.60E-34 |
| 1700042O10Rik | 19.07 | -7.17 | 1.21E-05 |
| Krt19 | 3429.17 | 7.15 | 0.00E+00 |
| Adgra1 | 7.03 | 7.15 | 1.43E-05 |
| Ccdc184 | 68.00 | 7.13 | 5.96E-16 |
| Plac8 | 734.06 | 7.13 | 1.70101E-43 |
| Irf2 | 74.07 | 7.13 | 7.83E-12 |
| Adh1 | 385.06 | 7.12 | 1.74E-119 |
| Lyve1 | 8.32 | 7.09 | 1.12E-05 |
| Kcnj5 | 9.64 | 7.09 | 1.01E-06 |
| 3830417A13Rik | 1189.41 | 7.09 | 3.12E-92 |
| Adamtsl3 | 39.03 | 7.05 | 2.21E-14 |
| Art4 | 37.56 | 7.05 | 3.40E-10 |
| Plau | 253.92 | 7.04 | 2.98E-36 |
| Frmd3 | 6.00 | 7.04 | 4.79E-04 |
| Hoxa11os | 8.35 | 7.02 | 2.94E-06 |
| Klk6 | 16.82 | 7.02 | 4.25E-07 |
| Sdsl | 47.67 | 7.02 | 3.39E-07 |
| Isx | 120.35 | 7.01 | 7.97E-08 |
| Hpcal4 | 68.14 | 7.00 | 8.19E-21 |

| | | | |
|---|---|---|---|
| Oit3 | 20.29 | 6.99 | 7.17E-07 |
| Nrp1 | 294.16 | 6.98 | 2.62E-83 |
| Trpc3 | 43.04 | 6.96 | 1.24E-11 |
| D7Ertd443e | 6.17 | 6.94 | 1.11E-05 |
| Hoxb6 | 32.71 | 6.94 | 4.90E-12 |
| Mir466f-2 | 10.02 | 6.93 | 9.79E-06 |
| Plppr4 | 14.35 | 6.92 | 2.10E-07 |
| Mycbpap | 110.10 | 6.90 | 2.75E-14 |
| Ifit3 | 9.93 | 6.90 | 6.43E-06 |
| Gm32591 | 12.27 | 6.90 | 5.05E-07 |
| Ankrd66 | 5.75 | 6.89 | 4.74E-05 |
| Ebf3 | 25.09 | 6.89 | 4.67E-09 |
| Tacstd2 | 482.67 | 6.88 | 3.39E-65 |
| Chrnd | 7.68 | -6.88 | 4.96E-06 |
| Fbxo32 | 396.78 | 6.87 | 9.50E-95 |
| Ifit1bl2 | 5.20 | 6.86 | 1.21E-04 |
| Ptprt | 57.57 | 6.85 | 2.35E-13 |
| Ankrd2 | 16.67 | 6.84 | 5.04E-07 |
| Abca4 | 425.94 | 6.83 | 2.73E-118 |
| Sct | 1334.87 | 6.82 | 1.36E-107 |
| Cryba4 | 16.59 | 6.82 | 2.34E-08 |
| Jpx | 26.94 | 6.81 | 2.31E-07 |
| Slco2b1 | 28.19 | 6.80 | 2.42E-07 |
| Ifi202b | 5.30 | 6.79 | 4.01514E-05 |
| Qrfpr | 21.00 | 6.79 | 5.56E-07 |
| Myl4 | 152.59 | 6.79 | 5.84E-18 |
| Gm10804 | 9.22 | 6.78 | 1.82E-05 |
| Tnf | 8.21 | 6.78 | 5.66121E-06 |
| Zfpm2 | 6.22 | -6.78 | 9.75E-03 |
| Scnn1b | 192.35 | 6.76 | 3.48E-45 |
| Maf | 55.34 | 6.74 | 3.21E-11 |
| Ido1 | 47.97 | 6.74 | 2.79E-12 |
| Prrx1 | 4.51 | 6.74 | 4.68E-04 |
| Insm1 | 8.02 | 6.73 | 3.17E-04 |
| Id2 | 872.91 | 6.72 | 1.43E-101 |
| Nrk | 5505.03 | 6.72 | 2.47E-94 |
| Il2rb | 265.16 | 6.70 | 1.85E-27 |
| Dct | 4.71 | 6.69 | 1.04E-03 |
| Arhgef6 | 837.87 | 6.69 | 2.4341E-69 |
| Cd82 | 1275.04 | 6.67 | 2.43E-28 |
| Cldn11 | 5.14 | 6.67 | 1.22E-04 |
| Gjb2 | 729.36 | 6.66 | 1.66E-24 |
| Ccdc194 | 6.33 | 6.66 | 6.47E-06 |
| Doxl2 | 12.94 | 6.65 | 1.20E-05 |
| Gsta1 | 117.38 | 6.65 | 4.61E-31 |
| Nkx2-3 | 5.58 | 6.65 | 5.02E-05 |
| Fam71f1 | 8.21 | 6.65 | 6.81E-05 |
| Vnn1 | 565.60 | 6.61 | 6.06E-60 |
| Ren1 | 70.99 | 6.59 | 9.09E-26 |
| Tnni1 | 113.06 | -6.59 | 1.32E-16 |
| Elf5 | 245.24 | 6.58 | 4.31E-56 |

Unbiased list of expression changes in iCdx2 ESCs from Day 0 to Day 7 of induction. Genes with an adjusted P-value < 0.01 shown, top 200 ordered by log fold change shown.

| Gene | Base mean | log2(FoldChange from D0 to D7) | Adjusted P-value |
|---|---|---|---|
| Prl2c5 | 148.85 | 12.68 | 1.18E-23 |
| Ctsll3 | 128.51 | 12.39 | 1.42E-21 |
| Prl7d1 | 119.37 | 12.35 | 6.16E-20 |
| Tlx2 | 246.65 | -12.02 | 1.54E-10 |
| Serpinb9g | 122.46 | 11.96 | 6.16E-21 |
| Fnd3c2 | 92.94 | 11.84 | 6.27E-19 |
| Psg29 | 83.41 | 11.79 | 1.70E-18 |
| Prl2a1 | 71.78 | 11.61 | 2.42E-03 |
| Olr1 | 343.67 | 11.54 | 1.82E-26 |
| Prl2b1 | 63.68 | 11.51 | 7.98E-19 |
| Prl7a1 | 125.62 | 11.48 | 7.55E-19 |
| Ctsr | 52.25 | 11.15 | 3.45E-03 |
| Itga2 | 58.62 | 11.14 | 1.16E-16 |
| BC051665 | 1518.17 | 11.11 | 4.91E-132 |
| Tmem54 | 60.96 | 11.11 | 1.84E-18 |
| Tpbpa | 49.80 | 11.06 | 1.99525E-17 |
| Pappa2 | 100.86 | 11.01 | 1.97E-18 |
| Vtcn1 | 111.64 | 10.95 | 3.86E-19 |
| 1600025M17Rik | 59.62 | 10.92 | 3.25826E-15 |
| Iigp1 | 86.62 | 10.83 | 2.56E-18 |
| C1ql2 | 60.21 | 10.69 | 1.47E-13 |
| Oasl2 | 68.67 | 10.63 | 2.04E-17 |
| Serpinb9c | 100.65 | 10.59 | 3.03E-17 |
| Entpd1 | 1317.85 | 10.47 | 4.74E-159 |
| Prl3d2 | 62.03 | 10.44 | 8.07E-07 |
| Flt3 | 40.89 | 10.40 | 5.03E-12 |
| Cdh5 | 1553.08 | 10.35 | 2.62E-139 |
| Postn | 44.45 | 10.27 | 2.06E-15 |
| Rhox12 | 39.77 | 10.27 | 2.66E-15 |
| Prl2c2 | 30.82 | 10.24 | 2.32E-11 |
| Gm4926 | 269.91 | 10.22 | 2.06E-40 |
| Prl4a1 | 26.85 | 10.19 | 8.66E-03 |
| Htr5a | 80.23 | -10.17 | 2.42E-09 |
| Scnn1g | 34.67 | 10.01 | 3.76E-12 |
| Fthl17a | 137.97 | 9.97 | 1.51E-20 |
| Fgd3 | 340.64 | 9.95 | 1.17E-53 |
| Brinp3 | 57.99 | -9.90 | 1.49E-09 |
| Ctsj | 696.22 | 9.86 | 7.55943E-74 |
| Prl3d1 | 150.72 | 9.84 | 4.22E-31 |
| Timd2 | 149.02 | 9.80 | 1.76E-15 |
| Slc6a2 | 954.30 | 9.72 | 6.94E-47 |
| Ctsq | 21.02 | 9.72 | 2.90E-04 |
| Wdr72 | 59.99 | 9.70 | 6.33E-15 |
| Serpinb9f | 90.65 | 9.69 | 3.45E-19 |
| Pparg | 337.77 | 9.67 | 3.54E-43 |
| Ifit1 | 28.84 | 9.67 | 4.20E-11 |
| Dppa1 | 739.76 | 9.59 | 2.75E-68 |
| Slc6a14 | 435.85 | 9.56 | 4.08E-85 |
| Ccrl2 | 30.03 | 9.55 | 1.27E-13 |
| Prl2c3 | 67.18 | 9.54 | 3.24E-16 |
| Igf2 | 6683.27 | 9.49 | 1.59E-42 |
| Serpinb9d | 22.75 | 9.47 | 4.91E-12 |
| Clic5 | 745.68 | 9.45 | 8.93E-57 |
| BC053393 | 698.18 | 9.43 | 7.11E-120 |
| Prl7b1 | 15.04 | 9.42 | 2.97E-08 |
| Grem2 | 61.80 | 9.41 | 3.27E-17 |
| 1810065E05Rik | 39.41 | 9.37 | 5.91E-13 |
| Epn3 | 554.62 | 9.37 | 1.97E-58 |
| Exoc3l4 | 626.68 | 9.30 | 1.47E-95 |
| Plau | 253.92 | 9.24 | 2.28E-60 |
| Ceacam15 | 13.93 | 9.24 | 5.77E-03 |
| Arhgef6 | 837.87 | 9.23 | 3.99E-121 |
| Il24 | 13.04 | 9.19 | 2.34E-08 |
| 1700086L19Rik | 25.33 | 9.18 | 7.27E-12 |
| Cxcr2 | 110.49 | 9.16 | 5.95E-25 |
| Gpr17 | 21.53 | 9.15 | 4.59E-12 |
| Serpinb9e | 287.66 | 9.14 | 7.93E-61 |
| Tmem132e | 182.70 | 9.14 | 7.58E-14 |
| Tacstd2 | 482.67 | 9.10 | 1.23E-107 |
| Mid2 | 83.75 | 9.08 | 1.26E-15 |
| Mir322 | 34.63 | 9.07 | 8.37E-13 |
| Prl5a1 | 13.55 | 9.06 | 1.54E-09 |
| Acta2 | 152.19 | 9.05 | 2.54E-22 |
| Procr | 381.37 | 9.05 | 6.72E-72 |
| Gbp2b | 35.28 | 9.04 | 4.03E-12 |
| Trpv5 | 226.27 | 9.04 | 2.05E-52 |
| Tmprss4 | 29.60 | 9.02 | 3.81E-12 |
| Dmkn | 397.10 | 9.02 | 4.30E-65 |
| Cts8 | 11.79 | 9.01 | 4.53E-08 |
| Pcdh12 | 570.34 | 9.00 | 2.77E-96 |
| Sqor | 57.06 | 8.97 | 3.29E-10 |
| Il2rb | 265.16 | 8.96 | 4.68E-44 |
| Hcls1 | 84.50 | 8.94 | 5.57E-22 |
| Gata3 | 352.57 | 8.93 | 3.44E-52 |
| Metrnl | 557.08 | 8.91 | 2.23E-98 |
| Lgals3bp | 623.28 | 8.89 | 5.26E-99 |
| Dhrs9 | 53.14 | 8.86 | 1.84E-14 |
| Mfsd6l | 18.82 | 8.83 | 3.21E-09 |
| Hand1 | 931.51 | 8.81 | 1.18E-149 |
| 1700071M16Rik | 31.70 | 8.80 | 3.82697E-10 |
| Slc16a2 | 75.04 | 8.80 | 3.57E-12 |
| Reg3b | 51.58 | 8.78 | 4.91E-12 |
| C3ar1 | 26.05 | 8.77 | 8.28E-10 |
| Doxl2 | 12.94 | 8.75 | 1.40E-08 |
| Plac1 | 513.80 | 8.69 | 3.20E-19 |
| Itgb6 | 46.42 | 8.67 | 1.39E-11 |
| Havcr1 | 171.92 | 8.64 | 1.61E-28 |
| Plac8 | 734.06 | 8.60 | 2.29E-58 |
| Dmrtc1a | 11.04 | 8.60 | 1.41E-06 |
| Psg22 | 10.09 | 8.60 | 3.89E-08 |
| Atp2c2 | 152.57 | 8.60 | 1.80E-23 |
| Gjb2 | 729.36 | 8.58 | 1.92E-33 |
| Pla2g4d | 49.99 | 8.56 | 5.13E-14 |
| Gbp3 | 22.40 | 8.54 | 1.04E-09 |
| Col3a1 | 30.39 | 8.53 | 1.37E-12 |
| Adgrf5 | 1881.06 | 8.53 | 1.72E-171 |
| Csf1r | 264.41 | 8.52 | 4.06E-60 |
| Gucy2c | 77.86 | 8.51 | 7.57E-15 |
| Mycbpap | 110.10 | 8.50 | 6.07E-21 |
| Tnfrsf9 | 252.02 | 8.47 | 6.77E-48 |
| Chst1 | 271.64 | 8.47 | 4.78E-19 |
| H19 | 47988.60 | 8.45 | 5.07E-48 |
| Mmp1a | 7.94 | 8.40 | 1.82E-06 |
| Hemgn | 13.19 | 8.40 | 1.88E-08 |
| AU021092 | 10.67 | 8.36 | 5.74E-09 |
| Rsad2 | 106.21 | 8.35 | 4.01E-28 |
| Mfsd2a | 346.99 | 8.35 | 1.61E-106 |
| Qrfpr | 21.00 | 8.35 | 9.14E-10 |
| Car4 | 496.34 | 8.34 | 5.27E-52 |
| Tns4 | 1966.05 | 8.33 | 1.35E-105 |
| Klk6 | 16.82 | 8.32 | 2.84E-09 |
| Syna | 51.28 | 8.31 | 7.93E-19 |
| Cyp11a1 | 103.29 | 8.31 | 7.95E-06 |
| Isg15 | 12.64 | 8.28 | 6.01E-07 |
| Tspan8 | 544.83 | 8.27 | 1.83E-92 |
| Mettl7b | 14.61 | 8.26 | 6.66E-09 |
| Vstm4 | 86.80 | 8.26 | 8.84E-14 |
| Krt18 | 14620.81 | 8.26 | 2.73E-296 |
| Hephl1 | 12.53 | 8.24 | 3.32E-08 |
| Chdh | 324.99 | 8.22 | 1.75E-60 |
| Nrk | 5505.03 | 8.22 | 1.31E-114 |
| Add2 | 412.09 | 8.22 | 1.67E-48 |
| Ceacam9 | 25.53 | 8.21 | 2.05E-11 |
| Enpp1 | 706.96 | 8.21 | 9.81E-106 |
| Mlkl | 14.49 | 8.20 | 1.53129E-07 |
| Duoxa2 | 52.15 | 8.19 | 1.32E-10 |
| Myo7b | 159.37 | 8.19 | 2.94E-25 |
| Cartpt | 18.24 | 8.19 | 1.63E-07 |
| Calr4 | 174.22 | -8.16 | 2.08E-15 |
| Vgf | 157.48 | 8.13 | 1.76E-46 |
| 3830417A13Rik | 1189.41 | 8.12 | 3.22E-103 |
| Spic | 250.50 | -8.11 | 9.73E-29 |
| Art4 | 37.56 | 8.10 | 5.73E-13 |
| Tnni1 | 113.06 | -8.10 | 2.21E-10 |
| Myl4 | 152.59 | 8.07 | 4.30E-22 |
| Mcub | 16.80 | 8.07 | 2.18E-09 |
| Nxf7 | 4846.84 | 8.06 | 4.10E-31 |
| Ctla2a | 6.84 | 8.06 | 1.67E-05 |
| Irf2 | 74.07 | 8.05 | 4.71E-14 |
| Dll4 | 65.00 | 8.03 | 5.24E-21 |
| Tll1 | 20.09 | -8.03 | 1.79E-06 |

| | | | |
|---|---|---|---|
| Elf5 | 245.24 | 8.02 | 2.24E-79 |
| Cd177 | 28.25 | 8.01 | 3.10E-11 |
| Arhgef37 | 109.11 | 7.98 | 9.11E-26 |
| Usp18 | 45.67 | 7.97 | 1.71E-16 |
| Thbd | 309.94 | 7.94 | 1.10E-97 |
| Igf2os | 501.33 | 7.93 | 5.57E-105 |
| Tmprss2 | 1813.23 | 7.92 | 7.56E-19 |
| Sod3 | 17.03 | 7.92 | 1.66E-08 |
| Scnn1b | 192.35 | 7.92 | 7.50E-59 |
| Cd82 | 1275.04 | 7.92 | 1.77E-32 |
| Noxa1 | 17.80 | 7.92 | 3.91E-09 |
| Jpx | 26.94 | 7.91 | 1.79E-09 |
| Entpd8 | 8.67 | 7.90 | 4.10E-07 |
| Prl6a1 | 18.06 | 7.86 | 2.18E-08 |
| Pde6h | 14.67 | 7.85 | 1.81E-07 |
| Adamts1 | 73.13 | 7.82 | 1.10E-15 |
| Fasl | 38.42 | 7.82 | 2.11E-13 |
| Pla2g7 | 1376.90 | 7.78 | 5.63E-86 |
| Sbsn | 217.13 | 7.77 | 2.13E-46 |
| Mir503 | 14.56 | 7.77 | 3.34E-08 |
| Krt19 | 3429.17 | 7.75 | 0.00E+00 |
| Pthlh | 21.89 | 7.75 | 5.32849E-09 |
| Slc24a3 | 47.75 | 7.74 | 3.07E-09 |
| Fam71f1 | 8.21 | 7.72 | 8.11E-06 |
| Tnf | 8.21 | 7.72 | 4.35E-07 |
| Rbbp8nl | 12.79 | 7.71 | 1.59362E-07 |
| Tnfaip2 | 956.90 | 7.70 | 4.23E-98 |
| Atp8b1 | 257.29 | 7.69 | 4.65E-57 |
| Cd93 | 7.38 | 7.69 | 1.16E-06 |
| Gm13003 | 6.81 | 7.68 | 1.30E-05 |
| Chrdl2 | 14.96 | 7.67 | 3.19E-08 |
| Pkp1 | 40.81 | 7.66 | 3.92E-11 |
| Gad2 | 17.64 | -7.65 | 8.57E-05 |
| Alpk1 | 100.21 | 7.64 | 5.35E-27 |
| Sct | 1334.87 | 7.64 | 7.98E-114 |
| Itga7 | 1314.14 | 7.62 | 1.31E-96 |
| Ifit3 | 9.93 | 7.62 | 1.30162E-06 |
| Prom2 | 44.65 | 7.61 | 6.64E-04 |
| Wnt7a | 6.36 | 7.60 | 4.32E-06 |
| Gm10804 | 9.22 | 7.55 | 3.84E-06 |
| Gm13580 | 73.47 | -7.55 | 1.12E-11 |
| Cdk5r2 | 22.61 | -7.52 | 3.03E-06 |
| Slc34a2 | 3213.65 | 7.52 | 1.09E-136 |
| Bgn | 93.48 | 7.52 | 1.90E-09 |
| Gsta1 | 117.38 | 7.51 | 3.76E-37 |
| Krt1 | 4.56 | 7.49 | 2.94E-04 |
| Irx1 | 102.52 | 7.48 | 5.69E-21 |
| 4930442L01Rik | 14.55 | 7.46 | 2.56E-07 |
| Nkx1-2 | 262.88 | -7.40 | 3.04E-05 |

Unbiased list of expression changes in iGata6 ESCs from Day 0 to Day 4 of induction. Genes with an adjusted P-value < 0.01 shown, top 200 ordered by log fold change shown.

| Gene | Base mean | log2(FoldChange from D0 to D4) | Adjusted P-value |
|---|---|---|---|
| Car4 | 496.34 | 11.56 | 6.49E-21 |
| BC025446 | 252.14 | 11.38 | 2.69E-20 |
| Gdf6 | 140.80 | 11.13 | 1.79E-18 |
| Habp2 | 494.28 | 11.04 | 6.43E-37 |
| Cubn | 18054.34 | 11.00 | 7.98E-195 |
| Slc34a2 | 3213.65 | 10.92 | 7.12E-140 |
| Col6a2 | 1638.58 | 10.91 | 1.62E-61 |
| Timd2 | 149.02 | 10.90 | 4.00E-19 |
| Tfec | 69.54 | 10.83 | 2.60E-16 |
| Lgals2 | 126.74 | 10.61 | 7.24E-18 |
| Muc13 | 518.80 | 10.56 | 1.37E-32 |
| Tlx2 | 246.65 | -10.53 | 2.83E-11 |
| Vcam1 | 120.80 | 10.46 | 6.55E-08 |
| Ttr | 55.19 | 10.12 | 1.43E-13 |
| Htr5a | 80.23 | -10.11 | 8.08E-13 |
| Pth1r | 3931.10 | 9.90 | 0 |
| Col6a1 | 3327.32 | 9.87 | 1.51E-142 |
| Slc24a3 | 47.75 | 9.81 | 1.04E-14 |
| Srgn | 5709.38 | 9.71 | 0 |
| Sel1l3 | 559.43 | 9.70 | 9.01E-29 |
| Serpina3i | 56.25 | -9.68 | 2.16E-13 |
| Ghr | 2061.59 | 9.67 | 6.51E-130 |
| Vstm4 | 86.80 | 9.63 | 1.42E-14 |
| Vsir | 4402.44 | 9.62 | 3.12E-156 |
| Il33 | 818.39 | 9.60 | 4.61E-76 |
| Myl3 | 408.44 | 9.54 | 2.32E-19 |
| Clic5 | 745.68 | 9.53 | 1.81E-52 |
| Ihh | 365.82 | 9.48 | 5.18E-21 |
| Fcgr2b | 53.92 | -9.46 | 7.89E-13 |
| Brinp3 | 57.99 | -9.43 | 2.84E-12 |
| S100g | 157.83 | 9.40 | 3.76E-27 |
| Col4a1 | 230646.30 | 9.37 | 1.11E-286 |
| Enpep | 1468.66 | 9.37 | 5.74E-125 |
| 1010001N08Rik | 169.72 | 9.34 | 1.55E-16 |
| Krt23 | 62.31 | 9.29 | 2.70E-13 |
| Col4a2 | 122006.40 | 9.27 | 5.02E-262 |
| Arhgef28 | 2705.31 | 9.22 | 7.09E-187 |
| Pdgfra | 10075.88 | 9.21 | 0 |
| Hkdc1 | 4455.71 | 9.21 | 5.31E-257 |
| Sp5 | 328.03 | -9.19 | 5.43E-10 |
| Foxq1 | 839.32 | 9.18 | 1.75E-16 |
| Lrrc32 | 82.24 | 9.18 | 3.89E-13 |
| Atp8a2 | 585.70 | 9.15 | 1.09E-41 |
| Fgg | 32.73 | -9.15 | 1.33E-07 |
| Nrn1 | 53.39 | 9.11 | 5.81E-13 |
| Dram1 | 946.30 | 9.08 | 1.27E-91 |
| Dclk3 | 74.78 | -9.07 | 5.14E-13 |
| Wnt9b | 52.24 | 8.99 | 3.56E-12 |
| Pik3ap1 | 377.07 | 8.96 | 1.20E-46 |
| Dab2 | 3286.32 | 8.95 | 8.00E-203 |
| P3h2 | 557.02 | 8.93 | 1.31E-62 |
| Zar1 | 20.22 | 8.90 | 1.59E-11 |
| Aff2 | 62.05 | 8.89 | 2.33E-06 |
| Dcstamp | 31.52 | -8.88 | 5.72E-11 |
| Crb1 | 19.29 | 8.88 | 2.73E-11 |
| Ptk6 | 1847.10 | 8.87 | 1.52E-124 |
| Nxf7 | 4846.84 | 8.87 | 5.40E-46 |
| Akr1c12 | 22.57 | 8.81 | 3.55E-10 |
| Mal | 420.94 | -8.81 | 7.06E-16 |
| Adora1 | 462.51 | 8.81 | 1.05E-59 |
| Plac8 | 734.06 | 8.79 | 1.39E-48 |
| Bmp2 | 406.00 | 8.75 | 3.26E-46 |
| Cyp26a1 | 391.45 | 8.73 | 4.79E-67 |
| Gsc | 107.52 | -8.70 | 4.45E-15 |
| Lgi4 | 33.08 | 8.70 | 1.40E-11 |
| H19 | 47988.60 | 8.63 | 2.10E-61 |
| Glipr1 | 833.88 | 8.63 | 4.54E-61 |
| Frzb | 1784.44 | 8.55 | 3.09E-128 |
| Slc39a8 | 437.11 | 8.55 | 3.10E-76 |
| Ctsh | 1061.62 | 8.54 | 1.19E-35 |
| Ugt2b34 | 846.80 | 8.52 | 1.39E-44 |
| Timp3 | 3866.32 | 8.48 | 1.88E-248 |
| Hnf1b | 488.90 | 8.43 | 3.94E-86 |
| Timp2 | 14714.37 | 8.40 | 1.51E-45 |
| Gadl1 | 61.27 | 8.39 | 8.24E-10 |
| Edn3 | 23.63 | 8.39 | 6.42E-11 |
| Trpm1 | 509.31 | -8.37 | 8.79E-39 |
| Slc6a11 | 17.80 | 8.34 | 3.42E-10 |
| Kctd16 | 31.99 | 8.34 | 1.05E-10 |
| Lrp2 | 46455.04 | 8.34 | 5.02E-208 |
| Slc6a2 | 954.30 | 8.31 | 2.17E-31 |
| Igf2 | 6683.27 | 8.30 | 8.34E-38 |
| Akr1c19 | 21.27 | 8.29 | 1.60E-08 |
| Foxa2 | 761.82 | 8.29 | 1.36E-95 |
| Ccdc33 | 54.13 | 8.27 | 8.58E-11 |
| Accsl | 889.11 | -8.27 | 6.90E-90 |
| Robo2 | 67.23 | 8.23 | 9.83E-12 |
| Nkx1-2 | 262.88 | -8.22 | 3.02E-07 |
| Hs3st1 | 1836.68 | 8.21 | 4.94E-198 |
| Plat | 4608.26 | 8.21 | 0 |
| Sox17 | 1603.50 | 8.20 | 3.43E-131 |
| Itgb6 | 46.42 | 8.18 | 1.24E-10 |
| Gm5535 | 18.57 | 8.15 | 5.98E-09 |
| Klb | 1578.66 | 8.14 | 7.22E-154 |
| BC016579 | 63.77 | 8.12 | 1.45E-10 |
| Ccdc180 | 17.85 | 8.11 | 1.03E-09 |
| Arsi | 574.81 | 8.10 | 2.08E-38 |
| Aqp1 | 20.98 | 8.09 | 1.57E-09 |
| Cdh6 | 200.21 | 8.05 | 2.87E-48 |
| Grrp1 | 74.45 | -8.03 | 1.48E-12 |
| Krt42 | 547.08 | -8.02 | 2.35E-62 |
| Fthl17a | 137.97 | 8.02 | 1.37E-10 |
| Mamdc2 | 54.33 | 8.00 | 2.07E-12 |
| Bmper | 883.55 | 7.97 | 1.46E-75 |
| Enpp2 | 75.82 | 7.97 | 7.11E-09 |
| 2300002M23Rik | 18.14 | -7.97 | 2.85E-08 |
| Nostrin | 719.58 | 7.94 | 1.75E-143 |
| Tll1 | 20.09 | -7.92 | 1.08E-08 |
| Defb42 | 17.32 | -7.92 | 6.58E-09 |
| Sox7 | 1675.40 | 7.92 | 5.19E-168 |
| Tpbgl | 424.59 | 7.91 | 1.20E-15 |
| Myo7a | 1362.83 | 7.91 | 9.76E-33 |
| Kcnq5 | 282.58 | 7.89 | 8.58E-31 |
| Amn | 1107.05 | 7.89 | 3.32E-149 |
| Mgat4c | 25.27 | -7.88 | 9.75E-09 |
| Kif12 | 10.48 | 7.87 | 6.91E-07 |
| Stpg1 | 35.50 | -7.84 | 1.00E-09 |
| Tns4 | 1966.05 | 7.84 | 3.80E-112 |
| Serpinb9e | 287.66 | 7.83 | 3.66E-10 |
| Gm805 | 17.06 | -7.82 | 3.24E-08 |
| Gm13889 | 59.68 | -7.82 | 6.56E-10 |
| Gm5868 | 9.75 | 7.81 | 5.48E-08 |
| Scn3a | 183.65 | -7.76 | 2.12E-28 |
| Dmkn | 397.10 | 7.76 | 9.35E-24 |
| Aqp8 | 2587.85 | 7.74 | 4.14E-27 |
| Emilin3 | 6397.03 | 7.74 | 1.10E-49 |
| Fxyd3 | 95.51 | 7.73 | 1.09E-12 |
| Pramel6 | 20.28 | -7.73 | 2.57E-08 |
| St6galnac5 | 20.18 | 7.72 | 1.43E-08 |
| Klhl34 | 39.17 | -7.72 | 4.28E-08 |
| Ramp1 | 40.05 | 7.71 | 6.76E-09 |
| Chst1 | 271.64 | 7.70 | 4.96E-21 |
| Kcnk13 | 12.32 | 7.70 | 1.66E-07 |
| Pcp4l1 | 36.90 | -7.70 | 2.00E-08 |
| Lamb1 | 106034.76 | 7.69 | 0 |
| Pdzd3 | 268.45 | 7.68 | 3.30E-66 |
| 5330411J11Rik | 18.86 | -7.61 | 2.01E-08 |
| Trpm2 | 222.99 | 7.60 | 1.86E-52 |
| Lypd3 | 18.02 | 7.58 | 1.86E-08 |
| Bace2 | 75.86 | 7.57 | 8.87E-16 |
| Lpar3 | 834.03 | 7.56 | 2.06E-102 |
| Trpc5os | 298.83 | 7.55 | 8.61E-46 |
| Kcna2 | 23.64 | -7.55 | 1.06E-07 |
| Ifi27l2b | 15.74 | 7.53 | 2.83E-08 |
| Olr1 | 343.67 | 7.52 | 6.33E-12 |
| Pou4f2 | 13.62 | -7.51 | 2.49E-06 |
| Ifitm7 | 13.62 | -7.49 | 1.38E-07 |
| 1600029O15Rik | 17.16 | 7.47 | 2.63E-07 |
| T | 45.70 | -7.46 | 8.44E-09 |
| Slc28a2 | 31.96 | 7.46 | 1.82E-08 |
| Apoa4 | 33.20 | 7.45 | 1.55E-09 |

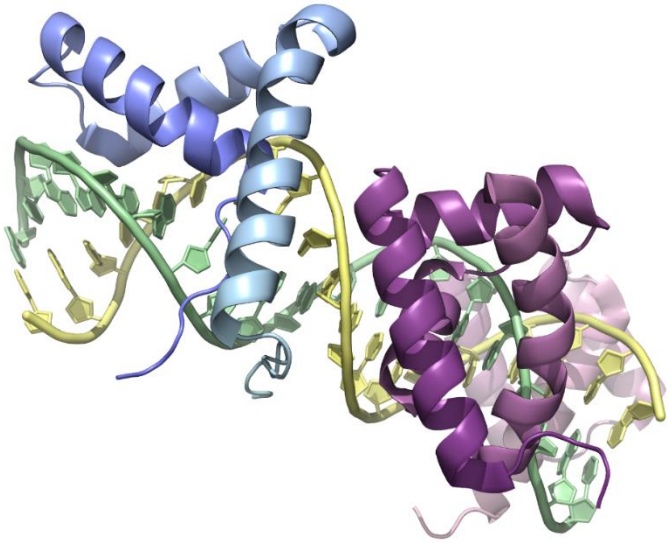
| | | | |
|---|---|---|---|
| Slc6a1 | 213.09 | -7.44 | 5.77E-19 |
| Acta1 | 120.10 | 7.44 | 1.01E-08 |
| Ppp1r3d | 31.56 | 7.40 | 1.13E-08 |
| Kcng1 | 372.95 | 7.39 | 2.85E-53 |
| Fgd3 | 340.64 | 7.38 | 3.61E-09 |
| Nrg2 | 1029.40 | 7.38 | 2.86E-179 |
| Cpm | 1884.99 | 7.36 | 1.00E-300 |
| Igf1 | 95.27 | 7.36 | 2.07E-07 |
| 4921513I03Rik | 20.83 | 7.34 | 5.85E-08 |
| Agtr1a | 63.24 | 7.33 | 1.27E-10 |
| Pyy | 15.81 | 7.32 | 9.99E-08 |
| Prokr2 | 23.15 | 7.32 | 5.53E-08 |
| Edil3 | 13.81 | -7.32 | 1.45E-07 |
| 1110002E22Rik | 22.51 | 7.31 | 2.49E-08 |
| Pcolce2 | 307.55 | -7.30 | 2.15E-55 |
| Cacna1d | 482.39 | 7.28 | 2.26E-86 |
| C2cd4a | 193.69 | 7.27 | 4.54E-56 |
| Nr2e3 | 9.74 | -7.27 | 3.13E-05 |
| Serpinb1b | 10.40 | -7.27 | 8.61E-05 |
| Thsd7b | 48.86 | 7.25 | 3.51E-09 |
| Il34 | 129.21 | -7.21 | 1.99E-18 |
| Dlx2 | 14.68 | 7.20 | 0.000331798 |
| Prame | 6.26 | 7.20 | 2.05E-05 |
| Ntrk3 | 181.54 | -7.20 | 2.22E-38 |
| Plac1 | 513.80 | 7.19 | 8.71E-12 |
| Mc5r | 10.64 | -7.18 | 0.00110486 |
| Mical2 | 3292.75 | 7.16 | 8.86E-162 |
| Egflam | 354.12 | 7.16 | 2.94E-47 |
| Foxf2 | 53.48 | 7.15 | 1.66E-18 |
| Frmpd1 | 1132.18 | 7.14 | 1.08E-81 |
| Rtn1 | 301.65 | 7.14 | 3.75E-51 |
| Myl9 | 594.20 | 7.13 | 3.13E-38 |
| Irf4 | 96.31 | 7.11 | 2.33E-13 |
| 4930591A17Rik | 26.58 | -7.10 | 3.52E-08 |
| Hcrtr2 | 10.01 | -7.10 | 9.68E-06 |
| Gm3629 | 15.58 | 7.07 | 5.71E-08 |
| Pcdhb13 | 5.51 | 7.06 | 0.000206034 |
| Lrrk1 | 402.38 | 7.05 | 1.11E-126 |
| Cend1 | 21.72 | 7.05 | 3.12E-07 |
| Arhgap6 | 828.78 | 7.04 | 5.98E-122 |
| Hsd17b14 | 244.88 | -7.03 | 2.69E-52 |
| Creb3l1 | 863.92 | 7.02 | 8.05E-79 |
| Cwh43 | 10.95 | -7.01 | 1.31E-05 |
| Apoa1 | 40.33 | 7.00 | 2.84E-10 |
| Maml3 | 832.68 | 6.99 | 1.09E-96 |
| Cdhr2 | 36.53 | 6.98 | 1.50E-12 |
| Steap4 | 11.82 | 6.96 | 2.63E-07 |
| Pof1b | 27.01 | -6.96 | 1.85E-07 |
| Cxcl10 | 8.84 | 6.96 | 2.93E-05 |

Unbiased list of expression changes in iGata6 ESCs from Day 0 to Day 7 of induction. Genes with an adjusted P-value < 0.01 shown, top 200 ordered by log fold change shown.

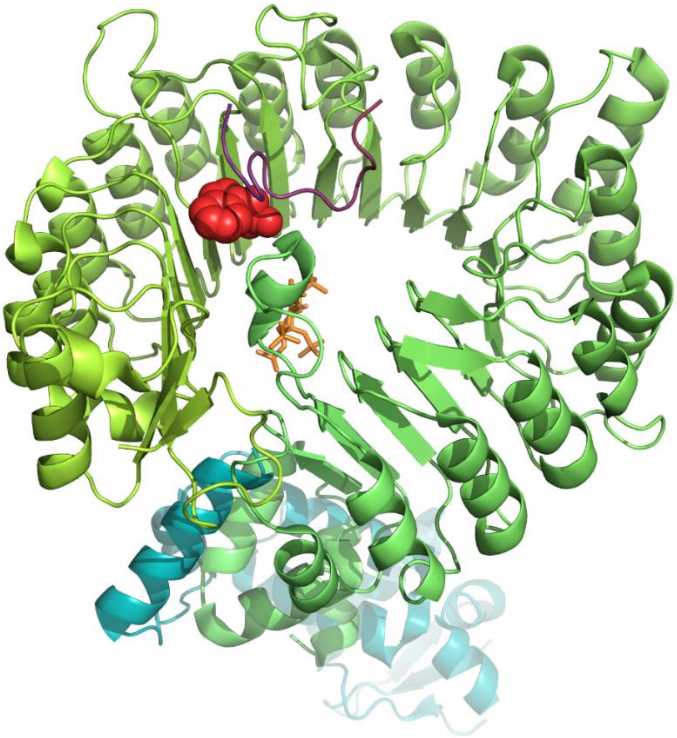
| Gene | Base mean | log2(FoldChange from D0 to D7) | Adjusted P-value |
|---|---|---|---|
| BC025446 | 252.14 | 12.45 | 5.54E-24 |
| Slc15a1 | 440.42 | -12.39 | 3.18E-15 |
| Accsl | 889.11 | -12.15 | 2.07E-17 |
| Myl3 | 408.44 | 11.71 | 1.37E-28 |
| Vcam1 | 120.80 | 11.40 | 3.22E-08 |
| Slc6a2 | 954.30 | 11.38 | 2.57E-56 |
| Slc34a2 | 3213.65 | 11.26 | 1.54E-146 |
| Myom2 | 174.66 | -11.26 | 2.78E-09 |
| Sel1l3 | 559.43 | 11.24 | 5.21E-38 |
| Scn3a | 183.65 | -11.20 | 3.14E-13 |
| Car4 | 496.34 | 11.12 | 2.35E-19 |
| Col6a2 | 1638.58 | 11.11 | 1.87E-59 |
| Ugt2b34 | 846.80 | 11.07 | 5.44E-72 |
| Kbtbd11 | 160.92 | -11.00 | 9.01E-13 |
| Gdf6 | 140.80 | 10.96 | 1.16E-17 |
| Gadl1 | 61.27 | 10.91 | 2.9103E-15 |
| Etv1 | 140.26 | -10.71 | 6.06E-11 |
| Arsi | 574.81 | 10.70 | 2.19E-65 |
| Mal | 420.94 | -10.66 | 2.3212E-12 |
| Vsir | 4402.44 | 10.61 | 4.74E-169 |
| Atp8a2 | 585.70 | 10.61 | 3.77E-55 |
| Lemd1 | 118.57 | -10.60 | 1.46E-11 |
| Timd2 | 149.02 | 10.60 | 4.03E-18 |
| Col6a1 | 3327.32 | 10.58 | 1.02E-147 |
| Ccdc33 | 54.13 | 10.57 | 8.52E-17 |
| Gsc | 107.52 | -10.55 | 6.39E-12 |
| Col4a1 | 230646.30 | 10.51 | 1.58E-289 |
| Ghr | 2061.59 | 10.47 | 1.76E-144 |
| Krt23 | 62.31 | 10.46 | 2.37E-16 |
| Frzb | 1784.44 | 10.45 | 2.10E-184 |
| Apod | 107.56 | -10.42 | 2.16E-11 |
| Col4a2 | 122006.40 | 10.35 | 1.41E-262 |
| Gkn1 | 35.96 | 10.34 | 8.36E-13 |
| Wnt9b | 52.24 | 10.30 | 2.20E-15 |
| Muc13 | 518.80 | 10.25 | 2.99E-30 |
| Padi4 | 88.36 | -10.23 | 3.19E-11 |
| Acta1 | 120.10 | 10.19 | 2.69E-15 |
| Tpbgl | 424.59 | 10.16 | 6.51396E-22 |
| Glipr1 | 833.88 | 10.15 | 5.06E-80 |
| Vstm4 | 86.80 | 10.13 | 7.06E-16 |
| Arhgef28 | 2705.31 | 10.10 | 4.89E-210 |
| Stra8 | 101.56 | -10.06 | 1.41E-07 |
| Wfdc1 | 58.20 | 10.04 | 2.23E-15 |
| Srgn | 5709.38 | 10.03 | 0.00E+00 |
| Timp2 | 14714.37 | 10.01 | 5.48E-52 |
| Itgb6 | 46.42 | 9.94 | 4.34E-15 |
| Hkdc1 | 4455.71 | 9.91 | 8.95E-268 |
| Ptk6 | 1847.10 | 9.91 | 2.42E-142 |
| Scn2b | 74.79 | -9.90 | 5.59E-09 |
| Spta1 | 66.26 | -9.85 | 2.53E-09 |
| Timm8a2 | 72.34 | -9.80 | 1.51E-10 |
| 1010001N08Rik | 169.72 | 9.80 | 1.23E-17 |
| Ankrd1 | 183.82 | 9.77 | 6.30E-33 |
| Pth1r | 3931.10 | 9.76 | 6.33E-305 |
| Bmp2 | 406.00 | 9.71 | 2.88E-56 |
| Chst1 | 271.64 | 9.71 | 1.14E-31 |
| Insyn1 | 114.21 | -9.71 | 3.62E-10 |
| Serpina3i | 56.25 | -9.69 | 1.30E-09 |
| Pcsk1 | 74.63 | -9.68 | 8.11E-10 |
| Cobl | 2006.99 | -9.68 | 1.60E-57 |
| Enpep | 1468.66 | 9.66 | 1.02E-126 |
| Slfn9 | 317.74 | -9.66 | 7.44E-16 |
| BC016579 | 63.77 | 9.65 | 2.02E-14 |
| Myl9 | 594.20 | 9.61 | 1.11E-63 |
| Fxyd3 | 95.51 | 9.52 | 1.49E-18 |
| Kctd16 | 31.99 | 9.49 | 2.42E-13 |
| P3h2 | 557.02 | 9.49 | 1.39E-69 |
| Lgi4 | 33.08 | 9.42 | 3.36E-13 |
| Platr17 | 55.47 | -9.40 | 3.74E-09 |
| Hs3st1 | 1836.68 | 9.39 | 3.31E-247 |
| Hsd17b14 | 244.88 | -9.39 | 3.96E-18 |
| Cdh2 | 84.38 | -9.38 | 2.62E-09 |
| Kcnq5 | 282.58 | 9.37 | 7.71E-41 |
| Dram1 | 946.30 | 9.32 | 2.16E-92 |
| Sema6d | 861.53 | 9.31 | 6.96E-136 |
| Pdgfra | 10075.88 | 9.29 | 0.00E+00 |
| Krt17 | 41.39 | -9.28 | 3.02E-08 |
| Slc28a2 | 31.96 | 9.28 | 2.63E-12 |
| Clic5 | 745.68 | 9.28 | 6.52E-49 |
| Tnfrsf8 | 53.51 | -9.26 | 1.43E-09 |
| Siglecf | 42.87 | -9.25 | 5.90E-09 |
| Smtnl1 | 44.09 | -9.22 | 9.91E-09 |
| Plac8 | 734.06 | 9.22 | 1.69E-50 |
| Lefty2 | 727.52 | -9.22 | 2.90E-16 |
| Lpar4 | 49.76 | -9.21 | 3.96E-07 |
| Emilin3 | 6397.03 | 9.21 | 6.59E-57 |
| Lipc | 41.74 | -9.21 | 3.02E-08 |
| Fam19a4 | 114.53 | -9.20 | 7.51E-10 |
| Adora1 | 462.51 | 9.20 | 7.94E-64 |
| 4921513I03Rik | 20.83 | 9.20 | 1.3153E-11 |
| Nrn1 | 53.39 | 9.20 | 4.75E-13 |
| Il34 | 129.21 | -9.20 | 2.67E-09 |
| Foxa2 | 761.82 | 9.14 | 1.31E-111 |
| Gm15462 | 54.49 | -9.13 | 8.67E-09 |
| Kcnj12 | 269.34 | -9.12 | 1.68E-16 |
| Sgo2b | 49.15 | -9.11 | 4.33E-08 |
| Gulp1 | 156.98 | -9.10 | 4.76E-09 |
| Aff2 | 62.05 | 9.09 | 5.72E-06 |
| Zscan5b | 46.06 | -9.08 | 5.46E-09 |
| Tmem169 | 19.05 | 9.08 | 1.30E-11 |
| Gsg1l | 18.52 | 9.05 | 8.22E-11 |
| Adm | 83.81 | 9.05 | 5.47E-13 |
| Ttr | 55.19 | 9.04 | 1.32E-10 |
| Vtcn1 | 111.64 | 9.03 | 1.69E-13 |
| Nkx1-2 | 262.88 | -9.02 | 2.43E-06 |
| Il33 | 818.39 | 8.94 | 7.28E-62 |
| Prdm14 | 718.48 | -8.93 | 1.26E-36 |
| Lrrc32 | 82.24 | 8.93 | 2.43E-12 |
| Halr1 | 50.94 | -8.91 | 9.97E-09 |
| Irf4 | 96.31 | 8.90 | 1.89E-19 |
| Dcstamp | 31.52 | -8.89 | 5.42E-08 |
| Acsbg2 | 37.11 | -8.88 | 1.25E-08 |
| Habp2 | 494.28 | 8.88 | 5.48E-24 |
| 1700029P11Rik | 48.02 | -8.87 | 1.86E-08 |
| Flt1 | 738.87 | 8.86 | 1.23E-72 |
| Nxf7 | 4846.84 | 8.86 | 1.20E-37 |
| Tns4 | 1966.05 | 8.84 | 7.84E-126 |
| Lpar3 | 834.03 | 8.83 | 1.93E-130 |
| Nlrp4a | 31.37 | -8.82 | 2.85E-07 |
| Sycp2l | 30.93 | -8.81 | 2.96E-07 |
| Gdf3 | 1014.14 | -8.79 | 6.34E-69 |
| Ramp1 | 40.05 | 8.78 | 6.07E-11 |
| 4930502E18Rik | 49.11 | -8.77 | 9.38E-08 |
| Kcng1 | 372.95 | 8.75 | 9.93E-71 |
| Timp3 | 3866.32 | 8.69 | 2.67E-229 |
| Rab38 | 159.80 | -8.68 | 2.20E-12 |
| Aqp1 | 20.98 | 8.68 | 1.65E-10 |
| Agtr1a | 63.24 | 8.67 | 4.14E-14 |
| St6galnac5 | 20.18 | 8.64 | 3.72E-10 |
| Tlx2 | 246.65 | -8.63 | 2.52E-09 |
| Spic | 250.50 | -8.63 | 9.29E-22 |
| Nanos3 | 31.93 | -8.62 | 1.60E-07 |
| Pik3ap1 | 377.07 | 8.60 | 2.97E-42 |
| Slc27a2 | 47.18 | -8.59 | 6.42E-08 |
| Plat | 4608.26 | 8.57 | 2.7214E-294 |
| Mical2 | 3292.75 | 8.57 | 9.71E-199 |
| Jakmip1 | 785.58 | -8.56 | 1.43E-49 |
| Mcf2 | 1734.74 | -8.56 | 3.69E-38 |
| Cdk14 | 252.16 | 8.55 | 3.29E-39 |
| Foxn4 | 35.34 | -8.55 | 3.53E-06 |
| Stmn3 | 166.51 | -8.54 | 1.38E-09 |
| 4930519F16Rik | 28.73 | -8.54 | 3.99E-08 |
| Dnajc5g | 32.25 | -8.53 | 1.73E-07 |
| Hdc | 261.97 | 8.52 | 9.77E-61 |
| Thbs1 | 433.16 | 8.52 | 2.60E-62 |
| Tnfaip2 | 956.90 | 8.51 | 7.69E-91 |
| 1600029O15Rik | 17.16 | 8.51 | 9.65E-09 |
| Prokr2 | 23.15 | 8.50 | 3.72E-10 |
| Trpm1 | 509.31 | -8.48 | 8.39E-29 |
| Ap3b2 | 429.44 | -8.45 | 2.44E-27 |
| 1700029M20Rik | 24.62 | -8.45 | 1.27E-07 |

| | | | |
|---|---|---|---|
| Sp5 | 328.03 | -8.44 | 1.26E-07 |
| Prmt8 | 27.38 | -8.43 | 1.73E-06 |
| Slfn10-ps | 28.21 | -8.43 | 7.42E-08 |
| Hesx1 | 31.11 | -8.43 | 2.73E-07 |
| Nanog | 12069.85 | -8.42 | 1.18E-102 |
| Ctgf | 4064.78 | 8.40 | 0.00E+00 |
| Bnc2 | 84.76 | -8.39 | 2.74E-08 |
| Tmem217 | 24.68 | -8.39 | 1.74E-06 |
| Notum | 284.36 | -8.36 | 5.17E-19 |
| Gm10863 | 26.54 | -8.35 | 1.92E-07 |
| 4930526L06Rik | 28.77 | -8.35 | 2.75E-07 |
| Dmgdh | 360.85 | -8.34 | 2.27E-24 |
| Krt42 | 547.08 | -8.31 | 4.43E-43 |
| Sox17 | 1603.50 | 8.30 | 5.09E-117 |
| Ccnd2 | 869.75 | 8.30 | 1.25E-24 |
| Tmem132e | 182.70 | 8.28 | 8.35E-23 |
| Thsd7b | 48.86 | 8.27 | 3.96E-11 |
| Trim54 | 22.44 | -8.25 | 7.08E-07 |
| Dlx2 | 14.68 | 8.25 | 1.03E-04 |
| Creb3l1 | 863.92 | 8.25 | 6.38E-95 |
| Calr4 | 174.22 | -8.24 | 5.67E-16 |
| Gpr88 | 10.93 | 8.24 | 1.19821E-05 |
| Myh7 | 78.00 | 8.24 | 7.84E-28 |
| Tmprss2 | 1813.23 | 8.23 | 4.55E-20 |
| Bche | 31.99 | -8.22 | 1.32E-07 |
| St8sia1 | 746.13 | -8.22 | 7.7929E-27 |
| Khdc3 | 401.58 | -8.21 | 9.52E-37 |
| Grid2 | 113.23 | -8.20 | 1.01E-12 |
| Arhgap6 | 828.78 | 8.20 | 3.52E-152 |
| Maml3 | 832.68 | 8.19 | 7.26E-129 |
| Calcoco2 | 1252.95 | -8.18 | 9.15E-69 |
| Robo2 | 67.23 | 8.16 | 4.58E-11 |
| Gm5535 | 18.57 | 8.15 | 1.38E-08 |
| Trh | 2226.84 | -8.15 | 6.51E-11 |
| Esrrb | 10284.46 | -8.15 | 3.63E-133 |
| Spata31d1a | 27.64 | -8.11 | 1.97E-07 |
| Slc24a3 | 47.75 | 8.11 | 3.31983E-10 |
| Aqp8 | 2587.85 | 8.10 | 3.58E-24 |
| 1700042O10Rik | 19.07 | -8.08 | 2.28E-05 |
| Osm | 156.71 | -8.07 | 1.31E-17 |
| Lamb1 | 106034.76 | 8.07 | 0.00E+00 |
| Cacna1d | 482.39 | 8.06 | 8.12E-100 |
| Ccdc180 | 17.85 | 8.06 | 2.39E-09 |
| Cacna1g | 71.12 | 8.06 | 9.30E-25 |
| Ihh | 365.82 | 8.06 | 1.42E-13 |
| Egfl6 | 35.75 | -8.05 | 6.74E-07 |
| Fcgr2b | 53.92 | -8.03 | 1.52E-07 |
| Brinp3 | 57.99 | -8.03 | 3.10E-07 |
| Bcas1 | 22.67 | -8.02 | 2.89E-07 |

# Appendix B



Oct1 (purple) and Sox2 (blue) bind to DNA (yellow and green), inducing bending. Structural data accessed from RCSB PDB public database entry 1O4X.
Williams, D.C., Cai, M., Clore, G.M., 2004. Molecular basis for synergistic transcriptional activation by Oct1 and Sox2 revealed from the solution structure of the 42-kDa Oct1.Sox2.Hoxb1-DNA ternary transcription factor complex. J. Biol. Chem. 279, 1449–57.



An AID-containing protein (IAA7, purple) interacts with Tir1 (green) via shared contact with auxin (red), leading to recruitment of Skp1 (cyan). Structural data accessed from RCSB PDB public database entry 2P1Q.
Tan, X., Calderon-Villalobos, L.I.A., Sharon, M., Zheng, C., Robinson, C. V, Estelle, M., Zheng, N., 2007. Mechanism of auxin perception by the TIR1 ubiquitin ligase. Nature 446, 640–645.

# Appendix C

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Genetics & Development

# Reprogramming human cells to naïve pluripotency: how close are we?

Lawrence E Bates and José CR Silva

Pluripotent stem cells (PSCs) have the potential to revolutionise biomedical science; however, while it is simple to reproducibly obtain comparable, stable cell lines in mouse, those produced from human material typically show significant variability both within and between cell lines. This is likely due to differences in the cell identity of conventional mouse and human PSCs. It is hoped that recently identified conditions to reprogram human cells to a naïve-like state will produce better PSCs resulting in reproducible experimental outcomes and more consistent differentiation protocols. In this review we discuss the latest literature on the discovery of human naïve-like stem cells and examine how similar they are to both mouse naïve cells and the preimplantation human epiblast.

**Address**
Wellcome Trust Medical Research Council Cambridge Stem Cell Institute and Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK

Corresponding author: Silva, José CR (jcs64@cam.ac.uk)

**Current Opinion in Genetics & Development** 2017, **46**:58–65

This review comes from a themed issue on **Cell reprogramming**

Edited by **Jianlong Wang** and **Miguel Esteban**

For a complete overview see the Issue and the Editorial

Available online 29th June 2017

**http://dx.doi.org/10.1016/j.gde.2017.06.009**



## Introduction

Studies in mouse embryonic stem cells (mESCs) over many years have led to a detailed understanding of this cell state. While mouse cells are typically grown in a state of naïve pluripotency, equivalent to the naïve epiblast of the preimplantation blastocyst [1], human cells are cultured in primed pluripotency conditions. These are more similar to the postimplantation epiblast where cells become primed for differentiation [2]. Consequently, there are significant difficulties in directly applying our knowledge from mouse ESCs to human systems.

There have been several attempts to generate human naïve pluripotent stem cells (nPSCs) over recent years. Most often when putative human naïve cells are generated *in vitro* they are analysed using criteria that are known to distinguish mouse naïve cells from primed cells. Such criteria include responses to extrinsic and intrinsic signalling pathways, the biophysical, biochemical and metabolic status of the cells, and the overall epigenetic and transcriptomic cell identity. However, recent advances in our understanding of the human embryo also allow direct comparisons to the naïve compartment *in vivo*. Recently, cells exhibiting human naïve epiblast molecular features have been described [3••,4••,5••]. Over the course of this review we shall examine how closely these match the state of both mouse naïve ESCs and what is known of the human blastocyst.

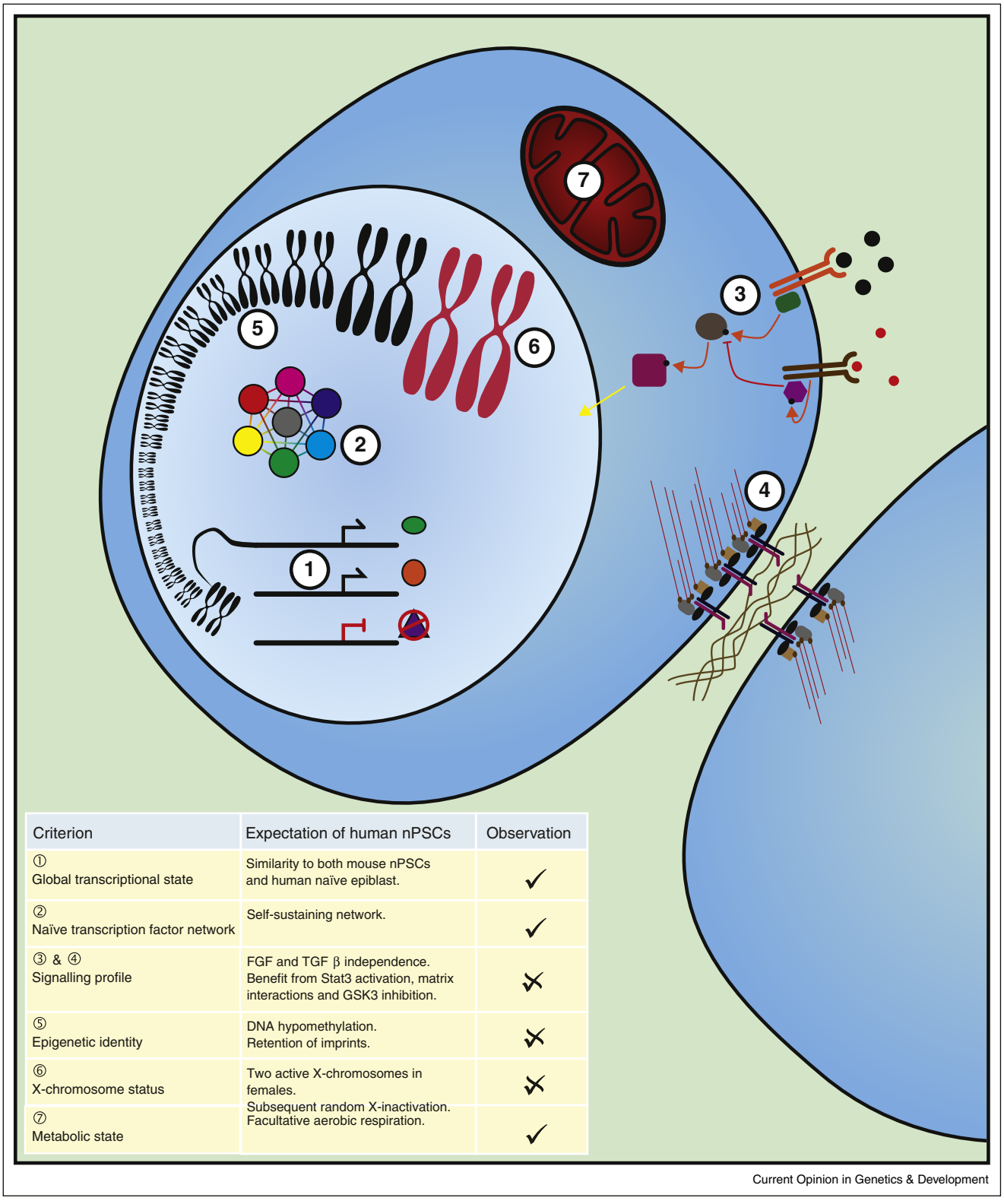
## Transcriptional identity

The transcriptional identity of a cell is often considered to be a readout of the cell's state (Figure 1, part 1). However, it is clear that the transcriptional state of cells is plastic, and a range of genes fluctuate in response to intracellular and extracellular conditions. Historically, it has not been practical to get a large number of high quality human embryos, necessary to obtain replicates of transcriptomic data with sufficient depth. Also, such embryos may be compromised as they have to be cultured *in vitro* in order to generate blastocyst-like embryos. Importantly, recent advances in RNA sequencing, particularly protocols for small cell numbers and even single cell sequencing, have made the analysis of these embryos possible.

Using such techniques, Yan *et al.* [6], and more recently Blakeley *et al.* [7], obtained single-cell transcript data from human embryos including late blastocyst stage embryos. Yan *et al.* observed four distinct cell types by unsupervised clustering which appear to represent two trophectoderm populations as well as extra-embryonic endoderm and epiblast cells based on their expression of known marker genes, as expected in a mature blastocyst. However, both studies identified only a handful of epiblast cells, giving a fairly small sample size for further analysis.

Comparing the human naïve induced pluripotent stem cells (hniPSCs) of Takashima *et al.*, the embryo derived human naïve ESCs (hnESCs) from Guo *et al.* [5], the embryonic naïve epiblast cells from Yan and from Blakeley, as well as mouse nESCs and conventional primed human ESCs, it is clear that the hniPSCs and embryo-derived hnESCs cluster closely together [5]. This indicates that they share a transcriptional identity. Notably, the human naïve-like cell lines cluster closer to mouse ESCs than to human blastocyst cells along a principle component of variation. It is possible that this separation

**Figure 1**




| Criterion | Expectation of human nPSCs | Observation |
|---|---|---|
| ① Global transcriptional state | Similarity to both mouse nPSCs and human naïve epiblast. | ✓ |
| ② Naïve transcription factor network | Self-sustaining network. | ✓ |
| ③ & ④ Signalling profile | FGF and TGF β independence. Benefit from Stat3 activation, matrix interactions and GSK3 inhibition. | ✗ |
| ⑤ Epigenetic identity | DNA hypomethylation. Retention of imprints. | ✗ |
| ⑥ X-chromosome status | Two active X-chromosomes in females. Subsequent random X-inactivation. | ✗ |
| ⑦ Metabolic state | Facultative aerobic respiration. | ✓ |

Current Opinion in Genetics & Development

Expected molecular signatures of human naïve pluripotent stem cells. A large number of processes control, and are influenced by, any cell state. Some of the factors that are particularly considered in this review are: **(1)** the transcriptional state of the cell. Functional components such as Oct4 and Nanog and marker genes such as Rex1 have been identified from mouse naïve cells and the human preimplantation epiblast, building a fingerprint of gene expression that should be present in naïve cells. **(2)** A core transcription factor network. The naïve state in mouse

is the result of generic adaptation of cells to *in vitro* culture. Interestingly, established human primed lines are separated from primed epiblast outgrowths along this same axis.

Theunissen *et al.* took a different approach to comparing their datasets to published human embryo data. They identified genes that are expressed in specific embryonic stages in the dataset of Yan *et al.* They then looked for the proportion of these genes that were differentially expressed between their hniPSCs and conventional primed hESCs. While genes specifically expressed in embryonic epiblast were enriched in the hniPSCs, so were genes specific to the morula and all other cell types of the late blastocyst [8•]. This is surprising, since unsupervised clustering of all single cells in the study by Yan *et al.* [6] indicated that these are well defined states, with epiblast cells segregating well from other cells of the blastocyst. As stated by Theunissen *et al.* [8•], this may indicate that their hniPSCs are in an earlier embryonic state, perhaps before full segregation of the inner cell mass. By analysing the expression of transposable elements, they found striking similarities between their cells and morula blastomeres. Interestingly, recent work from Petropoulos *et al.* [9•] suggests that the segregation of lineages occurs relatively late in human embryos, with co-expression of lineage-specific genes throughout the morula. The three lineages only begin to segregate on the onset of blastocyst formation.

## Transcription factor network

At the core of the naïve cell identity in mouse ESCs is a highly interconnected transcription factor network which shows remarkable redundancy [10,11] (Figure 1, part 2). Whereas the complete transcriptome may not be identical between cell lines and across passages, these factors are always expressed in naïve cells. They are also functionally conserved across a range of vertebrates in their ability to drive induction of naïve pluripotent stem cells [12–14]. Additionally, many of these factors are specific for nPSCs compared to primed post-implantation epiblast derived stem cells (EpiSCs) [2] making these a good subset of genes to use as markers of a naïve state. Takashima, Guo and Theunissen all investigated a panel of these genes and found that most were upregulated in their putative hniPSCs and hnESCs relative to primed cells [3••,4••,5••]. Neither *ESRRB* nor *KLF2* were upregulated in any of these naïve lines; however, this may be due to differences between primate and rodent and the redundant use of paralogue genes such as Klf4 [1,4••,7]. It is interesting, however, that both Takashima and Theunissen were able to efficiently induce a naïve-like state through exogenous expression of *KLF2* alongside *NANOG*. Takashima elegantly demonstrated that the behaviour of the transcription factor network in his hniPSCs closely corresponded to that of mouse ESCs with a knockout and rescue strategy. Mouse ESCs can support the single loss of either *Esrrb* or *Tfcp2l1* due to redundancies in the network [15,16], but it was expected that double knockout would lead to network collapse and subsequent differentiation [4••]. Accordingly, shRNA depletion of *TFCP2L1* in hniPSCs resulted in greatly reduced colony formation, indicating that most cells had stopped self-renewing. Application of exogenous *ESRRB* during this knockdown was able to rescue self-renewal. Together this provides strong evidence that an interactive transcription network highly similar to that in mouse is active in these cells.

## Exogenous ligands and intracellular signalling landscape

A broad array of signalling pathways interact to maintain or destabilise the naïve state in mouse ESCs (Figure 1, part 3). These cells are able to self-renew in the absence of external signals providing that certain pro-differentiation pathways are blocked, namely the MEK/ERK MAPK signalling axis and the GSK3β pathway which would otherwise destabilise the network in part through

**(Figure 1 Legend Continued)** has a self-sustaining network of transcription factors with many positive feedback loops to promote the maintenance of pluripotency. Notably, while many of these transcription factors are still expressed in primed cells, the network conformation is different, with factors binding to different enhancer elements and thus interacting in different ways. By exploring these interconnections, it is possible to test whether putative human naïve cells share the same connectivity and hence whether the network exists in a naïve configuration. **(3,4)** Environmental signals are key to maintaining cell states. In mouse, the naïve state can be maintained *in vitro* with LIF which activates downstream JAK/STAT signalling, an inhibitor of MEK/ERK signalling downstream of the FGF receptor, and an inhibitor of β-catenin degradation. The current human naïve culture conditions extend this with addition of a PKC inhibitor [4••], or BRAF, SRC and ROCK inhibitors [3••]. In addition to the response to ligands, cells interact physically with their neighbours and the extracellular matrix. Strong adherens junctions between cells provide the familiar dome-shaped morphology of naïve ESC colonies, and the ability to sense neighbours appears to be important for cell survival. **(5)** The epigenetic fingerprint of cells can be analysed in a similar manner to the transcriptional identity to build up a global picture of the cell state. Enhancer and promoter usage result in modification of histones and differential methylation of DNA. These profiles can be compared between cells. Additionally, the naïve state has additional epigenetic properties, such as global DNA hypomethylation and retention of imprinting marks which should be found in human naïve cells. **(6)** A key feature of the naïve state in female mouse cells is the presence of two active X-chromosomes. While the exact connection between naïve identity and X-chromosome status is still elusive, this is considered a hallmark of the naïve identity. While aspects of X-chromosome regulation differ between mouse and human, recent embryo work suggests that the human preimplantation epiblast shares this feature with mouse. **(7)** Many other elements of the cell are controlled by the cell state. One example is the switch between aerobic and anaerobic respiration. The naïve compartment of the embryo is considered to be facultatively aerobic, displaying relatively mature mitochondria, whereas other early embryonic cell states rely on anaerobic glycolysis for most of their energy requirements. While the cause of this switch is unknown, it is likely to be the result of integrating many other state-specific signals.

its role in the degradation of β-catenin [17]. However, these conditions are not optimal and extrinsic signals that enhance pluripotency and survivability are often added to the culture. Most common, is the LIFR/GP130 ligand LIF which activates the JAK/STAT pathway [18,19]. This is typically included with inhibitors of GSK3β and MEK to give chemically defined 2i LIF medium. This enhances the efficiency of induced naïve pluripotency [17,20] and maintains mouse naïve PSCs in a state closely resembling the naïve epiblast of the pre-implantation embryo [1,17,20].

Given the importance of LIF and downstream JAK/STAT signalling in reprogramming and maintenance of mouse naïve PSCs [21,22] and its ability to induce human cells with some naïve-like properties [23,24], it was surprising that Theunissen *et al.* [3••] found hLIF to be dispensable for culture of their hniPSCs. Indeed their microarray data (publically accessible on the GEO database, accession GSE59435) indicates that LIFR expression is fourfold lower in naïve cells than in their parental primed cells. In addition, Takashima *et al.* [4••] show that LIFR is expressed at a far lower level in hniPSCs than in mouse ESCs. While LIF signalling is a key feature of the mouse blastocyst and is important for survival in diapause [25], it is also known to be important in the process of implantation in both mouse and humans [26,27]. Further work will be required to test whether JAK/STAT signalling is indeed an important component of the human naïve state.

In mouse, Fgf2 and Activin A are both used to support the undifferentiated growth of primed EpiSCs [2,28,29], while Fgf signalling leads to differentiation of naïve ESCs and Activin signalling is dispensable [17,29–31]. While the current human naïve cells can be maintained in the absence of FGF and Activin [3••,4••], Theunissen *et al.* identified increased differentiation on their double inhibition [3••]. Notably in mouse, following activation by LIF stimulation, JAKs activate PI3K and the AKT signalling pathway [32]. This appears to have a role in ESC self-renewal, with cells treated with a PI3K or PDK1 inhibitor showing reduced proliferation and increased multi-lineage differentiation [33,34]. The PI3K/AKT pathway is known to be activated downstream of both FGF and Activin A/TGFβ signalling in various contexts [29,35,36], but most interestingly FGF2-induced AKT activation has been demonstrated in conventional primed human ESCs, where it is proposed to boost cell survival [37]; in this manner, FGF/Activin could be beneficial to the human naïve-like state in the absence of sufficient LIF signalling to induce the AKT signalling cascade.

Given the poor survival of hniPSCs Theunissen *et al.* [3••] found it necessary to maintain these with ROCK inhibitor. While ROCK inhibitor was found not to be strictly required by Takashima et al. [4••], it was used in combination with their t2iL+Gö in feeder-free culture and embryo derivation [4••,5••]. In mouse, poor survival on single-cell passaging is a more common trait in primed cells [2], which can be similarly rescued with application of ROCK inhibitor [38]. The precise manner in which ROCKi contributes to the enhancement of self-renewal is not clear. It appears that following single-cell dissociation, loss of focal adhesions between cells leads to activation of RHO/ROCK signalling. This results in enhanced actinomyosin contractility which induces apoptosis [39,40]. The improved proliferation observed while culturing the hniPSCs in the presence of ROCKi may suggest that they are not responding appropriately to cell-cell and cell-substrate contacts (Figure 1, part 4). In mouse, a switch from E-cadherin expression in naïve ESCs to N-cadherin in primed EpiSCs has been observed [41,42]. Examination of the microarray data from Theunissen *et al.* reveals a decrease in N-Cadherin on transition to the naïve state; however, both N-cadherin and E-cadherin are expressed in primed human ESCs, and there is no further increase in E-cadherin in the naïve cells [8•]. It could be interesting to examine other cell contact-sensing pathways such as YAP/TAZ signalling to see if they are compromised. Interestingly, it has been suggested that overexpression of YAP promotes the reprogramming of human primed cells to a naïve-like identity [43].

## Epigenetic fingerprint

Another distinctive feature of most cell states, and particularly the naïve state, is the epigenetic landscape (Figure 1, part 5). In mouse ESCs cultured in 2i LIF, there is a remarkable genome-wide reduction in DNA methylation which is also observed *in vivo* [44–46]. It has been unclear until recently whether this should also be expected in humans [47,48]. However, recent work suggests that DNA methylation shows the same trend in human as in mouse [49]. In line with this, Theunissen, Guo and Takashima all identified a decrease in global DNA methylation [4••,5••,8•], on the order of that observed in human embryos [49]. Beyond this global decrease, however, there are signs that DNA methylation may not be properly regulated in these cells. Imprinted loci are specifically marked by methylation on one of the parental chromosomes. Stable imprints are retained throughout development and can still be found in differentiated cells. Importantly, they are observed in primed human ESCs. On conversion to naïve conditions, however, many of these marks are lost [50•], which is an issue as this is linked to poor differentiation *in vitro* and shows links to developmental disorders and tumourogenisis [51].

Another epigenetic property of mouse naïve ESCs is the absence of a silent X-chromosome in females resulting in the presence of two active X-chromosomes (Figure 1, part 6). On fertilisation, the paternal X-chromosome is

specifically inactivated [52] and is only reactivated in the naïve epiblast at the blastocyst stage [52–55]. Shortly after implantation of the embryo a random X-chromosome is inactivated [56]. There has been some debate over the state in human blastocysts [57,58], but recent evidence shows that they do have two active X-chromosomes [9•]. In the early human female blastocyst, cells exhibit twice the amount of X-linked gene expression compared to counterpart male embryo cells [9•]. However, despite maintaining biallelic expression as the blastocyst develops further, transcription of X-linked genes is downregulated, a phenomenon termed 'dampening' of the X-chromosome [9•].

The status of the X-chromosome in female primed cells has also been somewhat contentious [59–61]. Primed cells have undergone X-inactivation, and there is no reactivation when reprogramming human somatic cells to primed iPSCs [61]. However, over prolonged culture of primed pluripotent cells, a phenomenon of erosion of X-inactivation can be observed [60–62]. It appears that expression of Xist, the master regulator of X-inactivation, becomes epigenetically silenced resulting in the subsequent activation of genes on the formerly inactive chromosome. This makes the presence of two active X chromosomes in primed human ESCs an epigenetic abnormality rather than a molecular feature of biological significance.

Recently, it was shown that in the reprogramming of human primed cells to a naïve-like state the silent X-chromosome reactivates [8•,63•]. Despite exhibiting biallelic expression, it was found that X-linked gene expression was not twice that of the cells of origin which contained a silent X chromosome [63•]. Instead it resembled the 'dampening' phenomenon observed *in vivo* in very late human blastocysts [8•,9•,63•]. Upon differentiation there was inactivation of the X chromosome. However, this was non-random and therefore not reflective of the process that occurs during development.

Together, these studies indicate that there are epigenetic differences between current human nPSCs, their *in vivo* counterpart, and mouse ESCs.

## Conclusion

By the majority of measures, the most up to date culture systems have produced human pluripotent cells with similarities to both mouse naïve ESCs and to the human preimplantation epiblast. Nonetheless, there are still significant discrepancies. The signalling pathways active in these cells and the transcription factor network they support, appear to be very similar to, yet far less stable than, their equivalents in mouse ESCs. It is currently not possible to say whether the reduced stability of the human naïve state is due to interspecies differences, suboptimal culture conditions, or the possibility that we have not yet isolated bona fide human nPSCs.

Evidence from Takashima and from Guo show that their naïve cells have undergone metabolic reprogramming, showing a significant level of mitochondrial respiration (Figure 1, part 7) [4••,5••]. This is typically associated with mature blastocysts, with cells before and after this stage relying more heavily on anaerobic respiration and displaying less mature mitochondria [64,65]. Additionally, naïve cells show increased glycolytic metabolism, inhibition of which appears to reduce their proliferation, demonstrating that the high level of metabolic activity is important to the maintenance of these cells [66]. On the other hand, some transcriptional data suggest that the culture systems may favour a less mature embryonic state. Finally, several issues such as reduced viability in single cell passaging and genomic instability could indicate suboptimal conditions for cell growth.

Interestingly, recent papers have identified novel hPSCs with broader chimerism and differentiation potential than naïve or primed cells. These respectively demonstrate the ability to form interspecies chimaeras and the ability to differentiate into both embryonic and extra-embryonic lineages *in vivo* [67,68]. The latter cells express naïve marker genes but show also an expression signature that is not similar to any embryonic cell type. Further characterisation of the novel human PSCs is now needed to ascertain their full properties and molecular signatures.

The next major hurdles in establishing hnPSCs as the standard for *in vitro* studies will require demonstrating superior differentiation potential and reliability compared to conventional human ES cultures, and methods to simplify the generation and culture of these cells. By identifying cell surface markers specific to hnPSCs, Collier *et al.* [69] present a step forward in facilitating the establishment of naïve cultures, as well as potentially allowing the study of reprogramming intermediates to dissect this interesting transition.

While the conditions for differentiation protocols may need to be optimised for these new cells, it will be important to learn whether the promises of more homogeneous, less cell-line dependent differentiation from a naïve starting population can be delivered. If so, then this cell state could take over to become the accepted standard starting point for drug discovery models, *in vitro* developmental studies, and possibly advances in cell therapies.

## Conflict of interest statement

Nothing declared.

## Acknowledgements

L.E.B. is the recipient of a MRC PhD studentship and J.C.R.S. of a Wellcome Trust Senior Research Fellowship (WT101861).

## References and recommended reading

Papers of particular interest, published within the period of review,

have been highlighted as:

- • of special interest
- •• of outstanding interest


1. Boroviak T, Loos R, Lombard P, Okahara J, Behr R, Sasaki E, Nichols J, Smith A, Bertone P: **Lineage-specific profiling delineates the emergence and progression of naive pluripotency in mammalian embryogenesis**. *Dev Cell* 2015, **35**:366-382.

2. Tesar PJ, Chenoweth JG, Brook FA, Davies TJ, Evans EP, Mack DL, Gardner RL, McKay RDG: **New cell lines from mouse epiblast share defining features with human embryonic stem cells**. *Nature* 2007, **448**:196-199.

3. •• Theunissen TW, Powell BE, Wang H, Mitalipova M, Faddah DA, Reddy J, Fan ZP, Maetzel D, Ganz K, Shi L *et al.*: **Systematic identification of culture conditions for induction and maintenance of naive human pluripotency**. *Cell Stem Cell* 2014, **15**:471-487.

This paper describes conditions (5i/L/A) that are sufficient to convert primed human ESCs to a naïve-like identity, and shows that these cells display several hallmarks of naïve pluripotency not seen in previous naïve-like culture systems.

4. •• Takashima Y, Guo G, Loos R, Nichols J, Ficz G, Krueger F, Oxley D, Santos F, Clarke J, Mansfield W *et al.*: **Resetting transcription factor control circuitry toward ground-state pluripotency in human**. *Cell* 2014, **158**:1254-1269.

This paper describes alternative conditions (t2iL+Gö) that can sustain a naïve-like identity similar to Theunissen *et al.* (2014) following transcription factor mediated reprogramming with similar distinctions over existing systems.

5. •• Guo G, von Meyenn F, Santos F, Chen Y, Reik W, Bertone P, Smith A, Nichols J: **Naive pluripotent stem cells derived directly from isolated cells of the human inner cell mass**. *Stem Cell Rep* 2016, **6**:437-446.

This paper demonstrates that cells can be derived from embryos directly into the conditions identified by Takashima *et al.* in [2], demonstrating that cells do not need to pass through a primed state or be subjected to transgene-induced reprogramming to reach a naïve-like state *in vitro*.

6. Yan L, Yang M, Guo H, Yang L, Wu J, Li R, Liu P, Lian Y, Zheng X, Yan J *et al.*: **Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells**. *Nat Struct Mol Biol* 2013, **20**:1131-1139.

7. Blakeley P, Fogarty NME, del Valle I, Wamaitha SE, Hu TX, Elder K, Snell P, Christie L, Robson P, Niakan KK: **Defining the three cell lineages of the human blastocyst by single-cell RNA-seq**. *Development* 2015, **142**:3151-3165.

8. • Theunissen TW, Friedli M, He Y, Planet E, O'Neil RC, Markoulaki S, Pontis J, Wang H, Iouranova A, Imbeault M *et al.*: **Molecular criteria for defining the naive human pluripotent state**. *Cell Stem Cell* 2016, **19**:502-515.

This paper further investigates the identity of cells cultured in the conditions identified by Theunissen *et al.* in [1] and Takashima *et al.* in [2], particularly comparing them to single-cell data from human embryos and primed cells.

9. • Petropoulos S, Edsgärd D, Reinius B, Deng Q, Panula SP, Codeluppi S, Plaza Reyes A, Linnarsson S, Sandberg R, Lanner F: **Single-cell RNA-Seq reveals lineage and X chromosome dynamics in human preimplantation embryos**. *Cell* 2016, **165**:1012-1026.

Using single-cell sequencing techniques, Petropoulos *et al.* demonstrate that cells of the pre-implantation human embryo maintain biallelic expression of X-linked genes, and that expression from the X-chromosome is dampened from E4 to E7, resulting in total expression from both chromosomes at a level similar to that seen in males.

10. Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J *et al.*: **Integration of external signaling pathways with the core transcriptional network in embryonic stem cells**. *Cell* 2008, **133**:1106-1117.

11. Yeo J-C, Ng H-H: **The transcriptional regulation of pluripotency**. *Cell Res* 2013, **23**:20-32.

12. Rodda DJ, Chew JL, Lim LH, Loh YH, Wang B, Ng HH, Robson P: **Transcriptional regulation of Nanog by OCT4 and SOX2**. *J Biol Chem* 2005, **280**:24731-24737.

13. Tapia N, Reinhardt P, Duemmler A, Wu G, Araúzo-Bravo MJ, Esch D, Greber B, Cojocaru V, Rascon CA, Tazaki A *et al.*: **Reprogramming to pluripotency is an ancient trait of vertebrate Oct4 and Pou2 proteins**. *Nat Commun* 2012, **3**:1279.

14. Theunissen TW, Costa Y, Radzisheuskaya A, van Oosten AL, Lavial F, Pain B, Castro LFC, Silva JCR: **Reprogramming capacity of Nanog is functionally conserved in vertebrates and resides in a unique homeodomain**. *Development* 2011, **138**:4853-4865.

15. Festuccia N, Osorno R, Halbritter F, Karwacki-Neisius V, Navarro P, Colby D, Wong F, Yates A, Tomlinson SR, Chambers I: **Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells**. *Cell Stem Cell* 2012, **11**:477-490.

16. Martello G, Bertone P, Smith A: **Identification of the missing pluripotency mediator downstream of leukaemia inhibitory factor**. *EMBO J* 2013, **32**:2561-2574.

17. Ying Q-L, Wray J, Nichols J, Batlle-Morera L, Doble B, Woodgett J, Cohen P, Smith A: **The ground state of embryonic stem cell self-renewal**. *Nature* 2008, **453**:519-523.

18. Williams RL, Hilton DJ, Pease S, Willson TA, Stewart CL, Gearing DP, Wagner EF, Metcalf D, Nicola NA, Gough NM: **Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells**. *Nature* 1988, **336**:684-687.

19. Smith AG, Heath JK, Donaldson DD, Wong GG, Moreau J, Stahl M, Rogers D: **Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides**. *Nature* 1988, **336**:688-690.

20. Silva J, Barrandon O, Nichols J, Kawaguchi J, Theunissen TW, Smith A: **Promotion of reprogramming to ground state pluripotency by signal inhibition**. *PLoS Biol* 2008, **6**:e253.

21. van Oosten AL, Costa Y, Smith A, Silva JCR: **JAK/STAT3 signalling is sufficient and dominant over antagonistic cues for the establishment of naive pluripotency**. *Nat Commun* 2012, **3**:817.

22. Burdon T, Stracey C, Chambers I, Nichols J, Smith A: **Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells**. *Dev Biol* 1999, **210**:30-43.

23. Chan YS, Göke J, Ng JH, Lu X, Gonzales KAU, Tan CP, Tng WQ, Hong ZZ, Lim YS, Ng HH: **Induction of a human pluripotent state with distinct regulatory circuitry that resembles preimplantation epiblast**. *Cell Stem Cell* 2013, **13**:663-675.

24. Chen H, Aksoy I, Gonnot F, Osteil P, Aubry M, Hamela C, Rognard C, Hochard A, Voisin S, Fontaine E *et al.*: **Reinforcement of STAT3 activity reprogrammes human embryonic stem cells to naive-like pluripotency**. *Nat Commun* 2015, **6**:7095.

25. Nichols J, Chambers I, Taga T, Smith A: **Physiological rationale for responsiveness of mouse embryonic stem cells to gp130 cytokines**. *Development* 2001, **128**:2333-2339.

26. Norwitz ER, Schust DJ, Fisher SJ: **Implantation and the survival of early pregnancy**. *N Engl J Med* 2001, **345**:1400-1408.

27. Zollner U, Bischofs S, Lalic I, Zollner K-P: **LIF and TNF alpha concentrations in embryo culture media are predictive for embryo implantation in IVF**. *Asian Pac J Reprod* 2012, **1**:277-282.

28. Greber B, Wu G, Bernemann C, Joo JY, Han DW, Ko K, Tapia N, Sabour D, Sterneckert J, Tesar P *et al.*: **Conserved and divergent roles of FGF signaling in mouse epiblast stem cells and human embryonic stem cells**. *Cell Stem Cell* 2010, **6**:215-226.

29. Lanner F, Rossant J: **The role of FGF/Erk signaling in pluripotent cells**. *Development* 2010, **137**:3351-3360.

30. Kunath T, Saba-El-Leil MK, Almousailleakh M, Wray J, Meloche S, Smith A: **FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from**

**self-renewal to lineage commitment**. *Development* 2007, **134**:2895-2902.

31. Fei T, Zhu S, Xia K, Zhang J, Li Z, Han J-DJ, Chen Y-G: **Smad2 mediates Activin/Nodal signaling in mesendoderm differentiation of mouse embryonic stem cells**. *Cell Res* 2010, **20**:1306-1318.

32. Hirai H, Karian P, Kikyo N: **Regulation of embryonic stem cell self-renewal and pluripotency by leukaemia inhibitory factor**. *Biochem J* 2011, **438**:11-23.

33. Welham MJ, Storm MP, Kingham E, Bone HK: **Phosphoinositide 3-kinases and regulation of embryonic stem cell fate**. *Biochem Soc Trans* 2007, **35**:225-228.

34. Paling NRD, Wheadon H, Bone HK, Welham MJ: **Regulation of embryonic stem cell self-renewal by phosphoinositide 3-kinase-dependent signaling**. *J Biol Chem* 2004, **279**:48063-48070.

35. Derynck R, Zhang YE: **Smad-dependent and Smad-independent pathways in TGF-beta family signalling**. *Nature* 2003, **425**:577-584.

36. Bauer J, Ozden O, Akagi N, Carroll T, Principe DR, Staudacher JJ, Spehlmann ME, Eckmann L, Grippo PJ, Jung B: **Activin and TGFβ use diverging mitogenic signaling in advanced colon cancer**. *Mol Cancer* 2015, **14**:182.

37. Eiselleova L, Matulka K, Kriz V, Kunova M, Schmidtova Z, Neradil J, Tichy B, Dvorakova D, Pospisilova S, Hampl A *et al.*: **A complex role for FGF-2 in self-renewal, survival, and adhesion of human embryonic stem cells**. *Stem Cells* 2009, **27**:1847-1857.

38. Ohtsuka S, Nishikawa-Torikai S, Niwa H: **E-cadherin promotes incorporation of mouse epiblast stem cells into normal development**. *PLoS ONE* 2012, **7**:e45220.

39. Ohgushi M, Matsumura M, Eiraku M, Murakami K, Aramaki T, Nishiyama A, Muguruma K, Nakano T, Suga H, Ueno M *et al.*: **Molecular pathway and cell state responsible for dissociation-induced apoptosis in human pluripotent stem cells**. *Cell Stem Cell* 2010, **7**:225-239.

40. Chen G, Hou Z, Gulbranson DR, Thomson JA: **Actin-myosin contractility is responsible for the reduced viability of dissociated human embryonic stem cells**. *Cell Stem Cell* 2010, **7**:240-248.

41. Bao S, Tang F, Li X, Hayashi K, Gillich A, Lao K, Surani MA: **Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells**. *Nature* 2009, **461**:1292-1295.

42. Weinberger L, Ayyash M, Novershtern N, Hanna JH: **Dynamic stem cell states: naive to primed pluripotency in rodents and humans**. *Nat Rev Mol Cell Biol* 2016, **17**:155-169.

43. Qin H, Hejna M, Liu Y, Percharde M, Wossidlo M, Blouin L, Durruthy-Durruthy J, Wong P, Qi Z, Yu J *et al.*: **YAP induces human naive pluripotency**. *Cell Rep* 2016, **14**:2301-2312.

44. Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, Regev A, Meissner A: **A unique regulatory phase of DNA methylation in the early mammalian embryo**. *Nature* 2012, **484**:339-344.

45. Reik W, Dean W, Walter J: **Epigenetic reprogramming in mammalian development**. *Science* 2001, **293**:1089-1093.

46. Leitch HG, McEwen KR, Turp A, Encheva V, Carroll T, Grabole N, Mansfield W, Nashun B, Knezovich JG, Smith A *et al.*: **Naive pluripotency is associated with global DNA hypomethylation**. *Nat Struct Mol Biol* 2013, **20**:311-316.

47. Fulka J, Fulka H, Slavik T, Okada K, Fulka J: **DNA methylation pattern in pig in vivo produced embryos**. *Histochem Cell Biol* 2006, **126**:213-217.

48. Dobbs KB, Rodriguez M, Sudano MJ, Ortega MS, Hansen PJ: **Dynamics of DNA methylation during early development of the preimplantation bovine embryo**. *PLOS ONE* 2013, **8**:e66230.

49. Guo H, Zhu P, Yan L, Li R, Hu B, Lian Y, Yan J, Ren X, Lin S, Li J *et al.*: **The DNA methylation landscape of human early embryos**. *Nature* 2014, **511**:606-610.

50. • Pastor WA, Chen D, Liu W, Kim R, Sahakyan A, Lukianchikov A, Plath K, Jacobsen SE, Clark AT: **Naive human pluripotent cells feature a methylation landscape devoid of blastocyst or germline memory**. *Cell Stem Cell* 2016, **18**:323-329.
This paper studies the epigenetic state of human naïve PSCs, both from blastocysts and the culture conditions identified in [1,2], in detail and identifies defects, particularly in the maintenance of methylation at imprinted loci.

51. Jelinic P, Shaw P: **Loss of imprinting and cancer**. *J Pathol* 2007, **211**:261-268.

52. Takagi N, Sasaki M: **Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse**. *Nature* 1975, **256**:640-642.

53. Sado T, Ferguson-Smith AC: **Imprinted X inactivation and reprogramming in the preimplantation mouse embryo**. *Hum Mol Genet* 2005, **14**:59-64.

54. Pasque V, Plath K: **X chromosome reactivation in reprogramming and in development**. *Curr Opin Cell Biol* 2015, **37**:75-83.

55. Silva J, Nichols J, Theunissen TW, Guo G, van Oosten AL, Barrandon O, Wray J, Yamanaka S, Chambers I, Smith A: **Nanog is the gateway to the pluripotent ground state**. *Cell* 2009, **138**:722-737.

56. Monk M, Harper MI: **Sequential X chromosome inactivation coupled with cellular differentiation in early mouse embryos**. *Nature* 1979, **281**:311-313.

57. van den Berg IM, Laven JSE, Stevens M, Jonkers I, Galjaard R-J, Gribnau J, van Doorninck JH: **X chromosome inactivation is initiated in human preimplantation embryos**. *Am J Hum Genet* 2009, **84**:771-779.

58. Okamoto I, Patrat C, Thépot D, Peynot N, Fauque P, Daniel N, Diabangouaya P, Wolf J-P, Renard J-P, Duranthon V *et al.*: **Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development**. *Nature* 2011, **472**:370-374.

59. Barakat TS, Ghazvini M, de Hoon B, Li T, Eussen B, Douben H, van der Linden R, van der Stap N, Boter M, Laven JS *et al.*: **Stable X chromosome reactivation in female human induced pluripotent stem cells**. *Stem Cell Rep* 2015, **4**:199-208.

60. Silva SS, Rowntree RK, Mekhoubad S, Lee JT: **X-chromosome inactivation and epigenetic fluidity in human embryonic stem cells**. *Proc Natl Acad Sci U S A* 2008, **105**:4820-4825.

61. Tchieu J, Kuoy E, Chin MH, Trinh H, Patterson M, Sherman SP, Aimiuwu O, Lindgren A, Hakimian S, Zack JA *et al.*: **Female human iPSCs retain an inactive X chromosome**. *Cell Stem Cell* 2010, **7**:329-342.

62. Mekhoubad S, Bock C, de Boer AS, Kiskinis E, Meissner A, Eggan K: **Erosion of dosage compensation impacts human iPSC disease modeling**. *Cell Stem Cell* 2012, **10**:595-609.

63. • Sahakyan A, Kim R, Chronis C, Sabri S, Bonora G, Theunissen TW, Kuoy E, Langerman J, Clark AT, Jaenisch R *et al.*: **Human naive pluripotent stem cells model X chromosome dampening and X inactivation**. *Cell Stem Cell* 2017, **20**:87-101.
This paper examines the status of the X-chromosome in human cells in great detail, looking specifically at pre-implantation blastocysts, primed cells in culture and naïve cells grown in the conditions identified in [1]. They find many features of X-chromosome status are similar between human niPSCs and the blastocyst, and show that reprogramming to the naïve state can reverse some abnormalities acquired in primed cell culture.

64. Zhou W, Choi M, Margineantu D, Margaretha L, Hesson J, Cavanaugh C, Blau CA, Horwitz MS, Hockenbery D, Ware C *et al.*: **HIF1α induced switch from bivalent to exclusively glycolytic metabolism during ESC-to-EpiSC/hESC transition**. *EMBO J* 2012, **31**:2103-2116.

65. Wilding M, Coppola G, Dale B, Di Matteo L: **Mitochondria and human preimplantation embryo development**. *Reproduction* 2009, **137**:619-624.

66. Gu W, Gaeta X, Sahakyan A, Chan AB, Hong CS, Kim R, Braas D, Plath K, Lowry WE, Christofk HR: **Glycolytic metabolism plays a**

**functional role in regulating human pluripotent stem cell state**. *Cell Stem Cell* 2016, **19**:476-490.

67. Wu J, Platero-Luengo A, Sakurai M, Sugawara A, Gil MA, Yamauchi T, Suzuki K, Bogliotti YS, Cuello C, Morales Valencia M *et al.*: **Interspecies chimerism with mammalian pluripotent stem cells**. *Cell* 2017, **168** 473–486.e15.

68. Yang Y, Liu B, Xu J, Wang J, Wu J, Shi C, Xu Y, Dong J, Wang C, Lai W *et al.*: **Derivation of pluripotent stem cells with in vivo**
ssue_no = "c" / >

**embryonic and extraembryonic potency**. *Cell* 2017, **169** 243–257.e25.

69. Collier AJ, Panula SP, Schell JP, Chovanec P, Plaza Reyes A, Petropoulos S, Corcoran AE, Walker R, Douagi I, Lanner F *et al.*: **Comprehensive cell surface protein profiling identifies specific markers of human naive and primed pluripotent states**. *Cell Stem Cell* 2017 http://dx.doi.org/10.1016/j.stem.2017.02.014.